



Db 943 nMetLysLysLysPheCysValLeuLysLysLysLysLeuSerGluAlaLysGluIleLysSe 963  
 QY 3169 ACAGTTAGAGAAACAAAGCTAAATGGGACAGAGCTCTCAGTGTGAGATGGCTTT 3228  
 Db 963 rGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuThrLe 983  
 QY 3229 AAATCAAGAAAGAGAGAGAGAAATGTCGATATATTAAAGAAAAAATTAGACCC-- 3286  
 Db 983 uAsnGlnGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1003  
 QY 3287 -----GAGAGCAACTTAGGAAAAAGTTAGAGTGAACACCACTTTGAACA 3333  
 Db 1003 uLeuGlyArgIleGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGluGlu 1023  
 QY 3334 GACTCTCAGAAATCAAGATATAGAAATGTAACAGTGAACAGTAAATTTGAATCAGTTTC 3393  
 Db 1023 nAlaLeuArgIleGlnAspIleGluLysLysSerValGluSerAsnLeuAsnGlnValSe 1043  
 QY 3394 TCACACTCATGAAGTGAATGATCTCTTCATGAAAAATTCATGTTGAAAAAGAAAT 3453  
 Db 1043 rHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGluIle 1063  
 QY 3454 TGCATGCTAAACTGGAAGTACCACTGAAACATCAACACAGGTGAAGGAAATAA 3513  
 Db 1063 eAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGlnLysGluAsnLys 1083  
 QY 3514 ATACTTTGAGGACATTAAAGATTTTACAGAAAGAGTCTGAACTTCAAAATGACCTAAA 3573  
 Db 1083 sTyrPheGluAspIleLysLysLysLysLysLysLysLysLysLysLysLysLysLys 1103  
 QY 3574 ACTGAACAGAAACAGTAAACAAAGGCGACTCTAGTATAGAGACGCTTAAAGTTCT 3633  
 Db 1103 sLeuLysGluLysSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysValLe 1123  
 QY 3634 GAGCGAGAGAACACAGTCTGACTTCTAAATGAAGGAA 3673  
 Db 1123 uIleAlaGluAsnThrMetLeuThrSerLysLysLysGlu 1136

RESULT 14

US-10-007-805-553  
 ; Sequence 553, Application US/10007805  
 ; Publication No. US20020150581A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Yugui  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Durham, Margarita  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
 ; FILE REFERENCE: 210121.470C10  
 ; CURRENT APPLICATION NUMBER: US/10/007,805  
 ; CURRENT FILING DATE: 2001-12-07  
 ; NUMBER OF SEQ ID NOS: 593  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 553  
 ; LENGTH: 1013  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-007-805-553

Alignment Scores:  
 Pred. No.: 2,96e-134 Length: 1013  
 Score: 2374.00 Matches: 512  
 Percent Similarity: 62.62% Conservative: 61  
 Best Local Similarity: 55.96% Mismatches: 126  
 Query Match: 36.82% Indels: 216

DB: 13 Gaps: 9  
 US-09-602-362E-26 (1-3673) x US-10-007-805-553 (1-1013)  
 QY 1199 TTGAGGCTCAAAAGAAACATCTGAGAAATTTTCATGGCCAGCAAAAGAAAGATCTAGG 1258  
 Db 12 MetSerProAlaLysGluThrSerGluLysPheThrTrpAlaLysGlyArgProArg 31  
 QY 1359 AAGATCATCGGAGGAGAAAGAAACATCTGTAAGACTGAATCCGTGGCAGAGTAACA 1318  
 Db 32 LysIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaAargValThr 51  
 QY 1319 CCTAATAAACTGAAGTTTGGAAAAAGAACATCTAATATGATGTCATCTCTCAAAA 1378  
 Db 52 SerAsnLysThrLysValLeuLysGlyArgSerLysMetIleAlaCysProThrLys 71  
 QY 1379 GAAACATCTACAAAAGCAAGTACAAATGTCGATGTGAGTTCTGTAGAGCCTATATCAGT 1438  
 Db 72 GluSerSerThrLysAlaSerAlaAsn----- 80  
 QY 1439 CTTTTCGCACACGCACTATTGAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAAT 1498  
 Db 80 ----- 80  
 QY 1499 CTTGCTACCAAGATTATCTCTAAGAGTGTGCAAGAAATTATACGTGTTTACCTGATGCT 1558  
 Db 80 ----- 80  
 QY 1559 ACATATCAAAAAGATATCAAAACAAATAAATCACAAAATAGAGATCAGATGTTCCCATCA 1618  
 Db 81 -----AspGlnArgPheProSer 86  
 QY 1619 GAATCCAAAACGAGGAGAGATGAAGAATATCTTGGGATTCGGGAGTCTCTTTGAGAGT 1678  
 Db 87 GluSerLysGlnGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSer 106  
 QY 1679 TCTGCAAGACTCAAGTGTCTATACCTGAGTCTATGATCAGAAAGTAAATGAGAGATAAT 1738  
 Db 107 SerAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsn 126  
 QY 1739 AGAGAGTAGAAGAGCTTCTCTGAGAAAGCCATCTCCCTTCAAGCCTCCGCTNGAAATGCAA 1798  
 Db 127 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 146  
 QY 1799 AAGACTGTTCCAAATAAAGCCTTTGAATTCAGAAATGAACAAACATGAGAGAGCTCAG 1858  
 Db 147 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 166  
 QY 1859 ATGTTCCCATCAGAATCCAAACAAAGAGCGATCAAGAAAAATCTTGGGATTCGAGAGT 1918  
 Db 167 MetPheProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 186  
 QY 1919 CCCTGTGAGACGGTTTCAAGAGAGTGTATTTACCCAAAGCTACACATCAAAAGAA 1978  
 Db 187 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 206  
 QY 1979 TTGATACCTTAAAGTGAATAATTAGAGAGTCTCTCTTAAAGATGGTCTCTTGAAGCCT 2038  
 Db 207 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 226  
 QY 2039 ACCTGTGGAAGAAAGTTCTCTTCCAAATAAAGCCTTAGAATTAAGACAGACAGAAACA 2098  
 Db 227 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 246  
 QY 2099 TTCAAAGCAGAGTCTCTGTATTAAGATGGTCTCTCTGAGCCTACCTGTGGAAGGAAGTT 2158  
 Db 247 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 266  
 QY 2159 TCTCTTCCAAATAAAGCCTTTAGAAATTAAGGACAGAGAAACACATCAAGCAGAGTCTCT 2218  
 Db 267 SerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAsp----- 284  
 QY 2219 GATAATGATGCTCTCTGAGAGCCTACTCTGTGGNAGAAAGTTCTCTTCCCAATAAAGCT 2278

Db 284 ----- 284  
QY 2279 TTAGAATTGACGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGATCCAAA 2338  
Db 285 ----- 285  
QY 2339 CAAAAGGATGATGAAGAAATCTTTGGGATTTTGAGAGTTTCCTTGAGACTCTCTTACAG 2398  
Db 293 GlnLysAspTyrGluGluAsnSerTrpAspThrGluSerLeuGluCysGluThrValSerGln 312  
QY 2399 AATGATGTGTTTCCCAAGCTCACATCAAAAAGAAATTCGATACCTTAAGTGGAAAA 2458  
Db 313 LysAspValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLys 332  
QY 2459 TTAGAAGAGTCTCCTGATAAAGATCGTCTCTGAAGCCTACCTGCGAATGAAATTTCT 2518  
Db 333 LeuGluGlySerProGlyLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 352  
QY 2519 CTTCCAAATTAAGCCCTTGAATTAAGGACAGAGAAACATTCAAAGCAGAGATGTGAGT 2578  
Db 353 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProGlu 372  
QY 2579 TCTGTAGAGTCCACATTCAGTCTTTTGGCAACCGACTACTGAAATTCACAGTCTACA 2638  
Db 373 Lys---ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal 387  
QY 2639 AAAAGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGAGCAACAAAGACAGACTAAGTGA 2698  
Db 388 -----ProAsnLysAlaLeuGluLeu 394  
QY 2699 CAACAGGAACGTGATATTGGCAATTATTGAACAGCTCCCAAGATCAAACAATAAGATG 2758  
Db 395 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro----- 407  
QY 2759 CCCACATCAGAATTAGGAGAGAAAGAGATACAAATCAACTTCAGATTCTGAGATTATC 2818  
Db 408 -----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 425  
QY 2819 TCTGTGAGTGATACACAGAAATATAGTGTCTTACCTGAGGCTACATATCAAAAAGAAATA 2878  
Db 426 CysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGlnLysGluIle 445  
QY 2879 AAGACAAACAAATGGCAATAGAGAGTCT----- 2908  
Db 446 AspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysSerPro 465  
QY 2908 ----- 2908  
Db 466 CysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 485  
QY 2909 -----CCTGAAAAGCCTTCTCACTTTGAGCTGCCACTGCAATGCAAACTCT 2956  
Db 486 LysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 505  
QY 2957 GTTCCAAATAAGGCTTAGAATGGAAGATAAACAACATTGAGAGCA----- 3004  
Db 506 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhe 525  
QY 3004 ----- 3004  
Db 526 ProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSerLeuArg 545  
QY 3004 ----- 3004  
Db 546 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 565  
QY 3005 -----GATTCAACTACCTATCAAAATCTTGGATCAGCTTCT 3043  
Db 566 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 585  
QY 3044 TCTTGTGAAGAGGAGGAACTTAAAGATACTGCAACAAATTCACACAAATG 3103  
|||||

Db 586 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 605  
QY 3104 GAACAATAAATAAGTTTCTGTACTACAAAGAACTGTCAAGGCGAAAGAAATA 3163  
Db 606 GluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 625  
QY 3164 AAATCACAGTTAGAGAACCAAAAGCTAAATGGGAACAAGAGCTCTGCAGTGTGAGATTG 3223  
Db 626 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnLeuCysSerValArgLeu 645  
QY 3224 CCTTAAATCAAGNAGAGAGAGAGAAATGTCGATATATATAAAGAAAAAATTAGA 3283  
Db 646 ThrLeuAsnGlnGluGluLysArgAsnAlaAspIleLeuAsnGluLysIleArg 665  
QY 3284 CCC-----GAAGAGCAACTTAGGAAAAAGTTAGAAAGTGAACCAACTT 3328  
Db 666 GluGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuValLysGlnLeu 685  
QY 3329 GAACAGACTCTCAGATACAGATATAGATTGAAAAGTGTAAACAGTAATTTGAATCAG 3388  
Db 686 GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln 705  
QY 3389 GTTCTCACACTCATGAAGTGAATGATCTCTTCATGAAATTCATGTTGAAAAAG 3448  
Db 706 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 725  
QY 3449 GAAATTCCTGCTAAACTGGAAGTAGCCACACTGAAACATCAACACAGGTGAGGAA 3508  
Db 726 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnLysGlu 745  
QY 3509 AATAAATACTTGAGGACATTAAAGATTTTACAAGAAAAAGAAATGCTGAACCTTCAATGACC 3568  
Db 746 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 765  
QY 3569 CTAAACTGAACAGAAAAACAGTAACAAAAAGGCGATCTCAGTATAGAGAGCGCTTAAA 3628  
Db 766 LeuLysLysLysGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 785  
QY 3629 GTTCTGACGACAGACACAGATGCTGACTCTTAATTTGAAGAA 3673  
Db 786 ValLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 800

## RESULT 15

US-10-076-622-553  
; Sequence 553 Application US/10076622  
; Publication No. US20030023036A1  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Sleath, Paul R.  
; APPLICANT: Persing, David H.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C11  
; CURRENT APPLICATION NUMBER: US/10/076,622  
; CURRENT FILING DATE: 2002-02-13  
; NUMBER OF SEQ ID NOS: 627  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 553  
; LENGTH: 1013  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-076-622-553

Alignment Scores:  
Pred. No.: 2,96e-134 Length: 1013  
Score: 2374.00 Matches: 512  
Percent Similarity: 62.62% Conservative: 61  
Best Local Similarity: 55.96% Mismatches: 126  
Query Match: 36.82% Indels: 216  
DB: 14 Gaps: 9

US-09-602-362E-26 (1-3673) x US-10-076-622-553 (1-1013)



```
Db      606 GluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 625
QY      3164 AAATCAGAGTTAGAGAACCAAAAGCTAAATCGGAACAAGAGCTCTGCGAGTGTGAGATTG 3223
Db      626 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeu 645
QY      3224 CCTTTAAATCAAGAAAGAGAGAGAAATGTCGATATATTAAAAAGAAAAAATTAGA 3283
Db      646 ThrLeuAsnGlnGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArg 665
QY      3284 CCC-----GAGAGCAACTTAGGAAAAAGTTAGAGTGAACACCAACTT 3328
Db      666 GluGluLeuGlyArgIleGluGlnHisArgLysGluLeuGluValLysGlnGlnLeu 685
QY      3329 GAACAGACTCTCAGAAATACAGATATAGAATTGAAAGTGTAACAAGTAATTTGAATCAG 3388
Db      686 GluGlnAlaLeuArgIleGlnAspIleGlnLeuLysSerValGluSerAsnLeuAsnGln 705
QY      3389 GTTCTCAGACTCATGAAAGTCAAAATGATCTCTTTCATGTGAAAATTGCATGTTGAAAAAG 3448
Db      706 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 725
QY      3449 GAAATGCCATCTAAAGTGAAGTAGCCACACTGAAACATCAACACAGGTGAAGGAA 3508
Db      726 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGlu 745
QY      3509 AATAAATACTTTGAGGACATTAAGATTTTACAAGAAAAAGAAATGCTGAACCTTCAAAATGACC 3568
Db      746 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 765
QY      3569 CTAAGACTGAACAGAAACAGTAAACAAGGGCATCTCAGTATAGAGAGAGCGCTTAAA 3628
Db      766 LeuLysLeuLysGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 785
QY      3629 GTTCTGACGGCAGAGAACACGATGCTGACTTCTAAATTGAAGGAA 3673
Db      786 ValLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 800
```

Search completed: July 15, 2004, 09:47:29  
Job time : 321.903 secs

1000

1000

1000

1000

1000

1000

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 15, 2004, 08:36:32 ; Search time 158.061 Seconds

(without alignments)  
13131.604 Million cell updates/sec

Title: US-09-602-362E-26

Perfect score: 6448

Sequence: 1 caagagcttgccgtacaga.....tgactcttaattgaaggaa 3673

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_n2p.model -DEV=xl  
-Q=/cgn2\_1/USPTO.spool\_p/US0960362/runat\_15072004\_093625\_21981/app\_query.fasta\_1.10325  
-DB=A Geneseq\_29Jan04 -QWTF=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LCOEPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MOD=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0960362 @CGN.1.1.729 @runat\_15072004\_093625\_21981 -NCPU=6 -ICPU=3  
-NC\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq\_29Jan04: \*  
1: geneseqp1980s: \*  
2: geneseqp1990s: \*  
3: geneseqp2000s: \*  
4: geneseqp2001s: \*  
5: geneseqp2002s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003bs: \*  
8: geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5938	92.1	1225	6 ABR47547	Abr47547 Breast ca
2	5173	80.2	1011	4 AAB84703	Aab84703 Amino aci
3	5173	80.2	1239	6 ABJ37789	Abj37789 Human tum
4	3458	53.6	1341	4 AAB84702	Aab84702 Amino aci
5	3458	53.6	1341	5 ABJ05537	Abj05537 Breast ca
6	3458	53.6	1341	6 ABJ37784	Abj37784 Human tum
7	3458	53.6	1341	6 ABR47548	Abr47548 Breast ca
8	3440	53.3	1349	6 ABJ37788	Abj37788 Human tum
9	2374	36.8	1013	6 ABJ37783	Abj37783 Human tum
10	2372.5	36.8	1095	4 AAU33357	Aau33357 Human bre

11	2372.5	36.8	1095	5 ABG78924	Abg78924 Human bre
12	2372.5	36.8	1095	6 ABJ37747	Abj37747 Human tum
13	2367	36.7	1002	4 AAU33351	Aau33351 Human bre
14	2367	36.7	1002	5 ABG78918	Abg78918 Human bre
15	2367	36.7	1002	6 ABJ37741	Abj37741 Human tum
16	2367	36.7	1002	8 ADE44437	Ade44437 Human bre
17	1776.5	27.6	661	6 ABJ37782	Abj37782 Human tum
18	1775	27.5	743	4 AAU33358	Aau33358 Human bre
19	1775	27.5	743	5 ABG78925	Abg78925 Human bre
20	1775	27.5	743	6 ABJ37748	Abj37748 Human tum
21	1769.5	27.4	650	4 AAB50263	Aab50263 Human bre
22	1769.5	27.4	650	4 AAG65983	Aag65983 B726P spl
23	1769.5	27.4	650	4 AAU33346	Aau33346 Human bre
24	1769.5	27.4	650	5 ABG78913	Abg78913 Human bre
25	1769.5	27.4	650	6 ABJ37736	Abj37736 Human tum
26	1769.5	27.4	650	8 ADE44421	Ade44421 Human bre
27	1542	23.9	512	4 AAB84701	Aab84701 Amino aci
28	1414.5	21.9	445	4 AAB50249	Aab50249 Human bre
29	1414.5	21.9	445	4 AAG65987	Aag65987 B726P spl
30	1414.5	21.9	445	4 AAU33350	Aau33350 Human bre
31	1414.5	21.9	445	5 ABG78917	Abg78917 Human bre
32	1414.5	21.9	445	6 ABJ37740	Abj37740 Human tum
33	1414.5	21.9	445	8 ADE44425	Ade44425 Human bre
34	1384	21.5	466	4 AAB50248	Aab50248 Human bre
35	1384	21.5	466	4 AAG65986	Aag65986 B726P spl
36	1384	21.5	466	4 AAU33349	Aau33349 Human bre
37	1384	21.5	466	5 ABG78916	Abg78916 Human bre
38	1384	21.5	466	6 ABJ37739	Abj37739 Human tum
39	1384	21.5	466	8 ADE44424	Ade44424 Human bre
40	1311	20.3	399	6 ABJ37785	Abj37785 Human tum
41	1298	20.1	399	6 ABJ37787	Abj37787 Human tum
42	1142.5	17.7	398	3 AAB07638	Aab07638 Amino aci
43	884.5	13.7	432	4 AAB50244	Aab50244 Human bre
44	884.5	13.7	432	4 AAG65982	Aag65982 B726P thi
45	884.5	13.7	432	4 AAU33345	Aau33345 Human bre

#### ALIGNMENTS

##### RESULT 1

ABR47547  
ID ABR47547 standard; protein; 1225 AA.  
XX  
AC ABR47547;  
XX  
XX  
DT 12-JUN-2003 (first entry)  
XX  
DE Breast cancer associated protein sequence SEQ ID NO:332.  
XX  
KW Human; breast cancer; cytostatic; gene therapy.  
XX  
OS Homo sapiens.  
XX  
XX WO2003004989-A2.  
XX  
PD 16-JAN-2003.  
XX  
PF 21-JUN-2002; 2002WO-US019669.  
XX  
PR 21-JUN-2001; 2001US-0299887P.  
PR 27-JUN-2001; 2001US-0301572P.  
PR 18-JUL-2001; 2001US-0306501P.  
PR 25-SEP-2001; 2001US-0325002P.  
PR 05-MAR-2002; 2002US-0362585P.  
PR 14-MAY-2002; 2002US-0380391P.  
XX (MILL-) MILLENIUM PHARM INC.  
XX  
PI Lillie J, Ganavarapu M, Glatt K, Hoersch S, Kamatkar S, Meyers RE;  
PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Mills GB;  
PI Bast RC, Hortobagyi GN, Pusztai L, Meric P, Sahin A, Mills GB;  
XX  
XX WPI; 2003-210381/20.  
DR

DR N-PSDB; ACC50245.  
XX  
PT Breast cancer diagnosis or treatment by comparing the level of expression  
PT of a marker in a patient sample with that in the control non-breast  
PT cancer sample.

XX Claim 1; SEQ ID NO 332; 128pp; English.

XX The present invention describes a method for assessing whether a patient  
XX is afflicted with breast cancer. The method comprises comparing the level  
XX of expression of a marker (gene/polypeptide) see ACC50076 to ACC50334 and  
XX ABR47386 to ABR47632) in a patient sample and the normal level of  
XX expression of the marker in a control non-breast cancer sample, where a  
XX significant increase in the level of expression of the marker in the  
XX patient sample and the normal level is an indication that the patient is  
XX afflicted with breast cancer. The breast cancer associated sequences from  
XX the present invention have cytostatic activities and can be used in gene  
XX therapy. The method is useful for diagnosing and treating breast cancer.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1225 AA;

Alignment Scores:  
Pred. No.: 0 Length: 1225  
Score: 5938.00 Matches: 1159  
Percent Similarity: 99.57% Conservative: 1  
Best Local Similarity: 99.48% Mismatches: 4  
Query Match: 92.09% Indels: 1  
DB: 6 Gaps: 0

US-09-602-362E-26 (1-3673) x ABR47547 (1-1225)

QY 181 ATGAGAGGCTCTAGCTCCGCTGGCAAGGCGTGGGGGCGCGGCGCCCGGCAACCC 240  
DB 1 MetLysArgLeuLeuAlaAlaGlyLysGlyValArgGlyProGluProProAsnPro 20  
QY 241 TTCAGGACGGGTCTACACTGAGAGGACTACGGGACCATCTACTCTGGGACTAGG 300  
DB 21 PheSerGluArgValThrThrGluLysAspThrGlyThrIleTyPheGlyAspLeuGly 40  
QY 301 AAGATCCATACAGCTGCTCCCGGGGCCAAGTCCAGAAGCTGGAGAAGATCACAGTAGG 360  
DB 41 LysIleHisThrAlaAlaSerArgGlyGlnValGlnLysLeuGluLysMetThrValGly 60  
QY 361 AAGAGCCCGCTCACTGAAACAAAGAGATATGAGAGAGAGGACTGCTACACTGGGCC 420  
DB 61 LysLysProValAsnLeuAsnLysArgAspMetLysLysArgThrAlaLeuHisThrAla 80  
QY 421 TGTGTCAATGGCCATGCANAGTAGTAACATTTCTGTAGACAGAAAGTGCNGCTTAAT 480  
DB 81 CysValAsnGlyHisAlaGluValThrPheLeuValAspArgLysCysGlnLeuAsn 100  
QY 481 GTCCTTGATGGCAAGGAGGACACCTCTGATGAGAGCTCTACAAATGCGAGGGAAGCT 540  
DB 101 ValLeuAspGlyGluGlyArgThrProLeuMetLysAlaLeuGlnCysGluArgGluAla 120  
QY 541 TTGTGCAATATTTCTCATAGATGCTGGTCTCATCTTAAATATGATGATGTATGGCAA 600  
DB 121 -CysAlaAsnIleLeuIleAspAlaGlyAlaAspLeuAsnThrValAspValThrGlyAs 140  
QY 601 CAGGCTCTCCATTATGCGGTTTATAGTGAGAAATTTAATATGTTGGCAACACTGCTCTC 660  
DB 140 nThrAlaLeuHisThrAlaValThrSerGluAsnLeuLeuMetValAlaThrLeuLeuSe 160  
QY 661 CTATGGTCAGTCATCGAGGCTCAACACAGGCTAGCTCACACCCCTTTTACTGGCCAT 720  
DB 160 rTyGlyAlaValIleGluValGlnAsnLysAlaSerLeuThrProLeuLeuLeuAlaI 180  
QY 721 ACAGAAAGAACGACCAACTGTGGAATTTTACTAAACAAAAATGCAAAATGCAACGC 780  
DB 180 eGlnLysArgSerLysGlnThrValGluPheLeuLeuThrLysAsnAlaAlaAsnAla 200

QY 781 ATTAATGAGCTAAATGACAGCCCTCATGCTTGCATATGTGAAGGCTCATCAGAT 840  
DB 200 aPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIleCysGluGlySerSerGluI 220  
QY 841 AGTCGGCATCTTCTTCAGCAAAATGTTGACGCTTTTGTGTGAGACATACATGATAC 900  
DB 220 eValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaGluAspIleHisGlyIle 240  
QY 901 TGCAAGAACGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 240 rAlaGluArgTyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 260  
QY 961 TATACGAAATTAACCTTAAATAATCCTCAAAATCAATCAAGAGGAAACATCTACAGGA 1020  
DB 260 sIleArgLysLeuProLysAsnProGlnAsnThrAsnProGluGlyThrSerThrGly 280  
QY 1021 ACCTGATGAGGCTGCACCTTGGCGGAAAGAACACCTGACACGCGTGAAGCTTGTCTGA 1080  
DB 280 rProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGluSerLeuLeu 300  
QY 1081 AAAAACACCTGACGAGGCTGCACGCTTGTGTGAGGAAAGCTGTGCCAAAATCAATGCT 1140  
DB 300 uLysThrProAspGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 320  
QY 1141 GGGGAAAGCAACATCTGGAAGGTTTGAACAGTCAACAGAGAAACACCTAGGAAATTTT 1200  
DB 320 uGlyLysAlaThrSerGlyLysPheGluGlnSerThrGluGluThrProArgLysIle 340  
QY 1201 GAGGCTTACAAAAGAAACATCTGAGAAATTTTCATGCCCAAGAAAGAAAGTCTAGGAA 1260  
DB 340 uArgProThrLysGluThrSerGluLysPheSerTrpProAlaLysGluArgSerArg 360  
QY 1261 GATCACATGGAGGAAAGAAACATCTGTAAGACTGATCGGTGGCAGGATACACC 1320  
DB 360 sIleThrTrpGluGluLysGluThrSerValLysThrGluCysValAlaGlyValThr 380  
QY 1321 TAATAAACCTGAAAGTTTGGAAAAAGAACATCTAATATGATTGATGCTCTACAAAGA 1380  
DB 380 oAsnLysThrGluValLeuGluLysGlyThrSerAsnMetIleAlaCysProThrLys 400  
QY 1381 AACATCTACAAAAGCAAGTACAAATGTGGATGTGAGTGTCTGTAGACCTATATTCAGT 1440  
DB 400 uThrSerThrLysAlaSerThrAsnValAspValSerSerValGluProIlePheSer 420  
QY 1441 TTTTGGCACACGACATTTGAAAATTCACAGTGTACAAAAGTTGAGCAAGACTTTAATCT 1500  
DB 420 uPheGlyThrArgThrIleGluAsnSerGlnCysThrLysValGluGluAspPheAsn 440  
QY 1501 TGCTACCAAGATTTCTTAAGAGTCTGTCACAGAAATTATAGTGTGTACCTGATGCTAC 1560  
DB 440 uAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyThrCysLeuProAspAla 460  
QY 1561 ATATCAAAAAGATATCAAAACATTAATCAAAAATAGAGATCAGATGTTCCCATCAGA 1620  
DB 460 rTyGlnLysAspIleLysThrIleAsnHisIleGluAspGlnMetPheProSerG 480  
QY 1621 ATCAAAACGAGAGGAGAGTGAAGAAATTTCTGGGATTTCTGGGAGTCTCTTTGAGAGT 1680  
DB 480 uSerLysArgGluGluAspGluGluTySerTrpAspSerGlySerLeuPheGluSerSe 500  
QY 1681 TGCAAGACTCAAGTGTGTATACCTGAGTCTATGATCAGAAAGTAAATGAGATAAATAG 1740  
DB 500 rAlaLysThrGlnValCysIleProGluSerMetTyGlnLysValMetGluIleAsn 520  
QY 1741 AGAAGTAGAAGAGCTCTCTGAGAAAGCCTCTGCTTCAAGCCCTGCGGTGAAATGCAAAA 1800  
DB 520 gGluValGluGluLeuProGluLysProSerAlaPheLysProAlaIleGluMetGln 540  
QY 1801 GACTGTTCCAAATTAAGCCCTTTGAATTTGAAGATGAACAAACATTTGAGAGCAGCTCAG 1860  
DB 540 sThrValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAlaGlnMe 560



1861 GTTCCATCAGATCCAAACAAAGGACGATGAAGAAAATCTTGGGATTCGAGAGTCC 1920  
Db tPheProSerGluSerGlyGlnLysAspAspGluGluAsnSerTrpAspSerGluSerPr 580  
1921 CTGTGAGACGGTTTCACAGAGAGTGTATTTACCCAAAGCTACACATCAAAAGAAAT 1980  
Db oCysGluThrValSerGlnLysAspValTrpLeuProLysAlaThrHisGlnLysGluPh 600  
1981 CGATACCTTAAGTGGAAAATAGAGAGTCTCTGTTAAAGATGGTCTTCTGAGAGCTAC 2040  
Db eAspThrLeuSerGlyLysLeuGluGluSerProValLysAspGlyLeuLeuLysProTh 620  
2041 CTGTGGAAGGAAAGTTCTCTTCCAAATAAAGCCCTTAGAATTAAGGACAGAGAAACATT 2100  
Db rCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPh 640  
2101 CAAGCAGAGTCTCTGTATAAGATGCTCTCTGAAGCTACCTGTCGAAGGAAGTTTC 2160  
Db eLysAlaGluSerProAspLysAspGlyLeuLeuLysProThrCysGlyArgLysValSe 660  
2161 TCTTCCAAATAAAGCCTTAGAATTAAGGACAGAGAAACACTCAAGCAGAGTCTCTCGA 2220  
Db rLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrLeuLysAlaGluSerProAs 680  
2221 TAATGATGGTCTCTGAAGCCTACCTGTGGAAGGAAGTTCTCTTCCAAATAAAGCTTT 2280  
Db pAsnAspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLe 700  
2281 AGAATTGAAGACAGAGAAACATTCAAGCAGCTCAGATGTTCCATCAGAAATCCAAACA 2340  
Db uGluLeuLysAspArgGluThrPheLysAlaAlaGlnMetPheProSerGluSerLysGl 720  
2341 AAAGATGATGAAGAAAATCTCTGGGATTTGAGAGTTTCTCTGAGACTCTCTTACAGAA 2400  
Db nLysAspAspGluGluAsnSerTrpAspPheGluSerPheLeuGluThrLeuLeuGlnAs 740  
2401 TGATGTGTTTACCCAGGCTACACATCAAAAGAAATTCGATACCTTAAGTGAAGAATT 2460  
Db nAspValCysLeuProLysAlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLe 760  
2461 AGAAGATCTCTGTATAAGATGGTCTCTCAAGCCTACCTGCGAATGAAAATTTCTCT 2520  
Db uGluGluSerProAspLysAspGlyLeuLeuLysProThrCysGlyMetLysLeSerLe 780  
2521 TCCAAATAAAGCCTTAGAATTAAGGACAGAGAAACATTCAAGCAGAGGATGTGAGTTC 2580  
Db uProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaGluAspValSerSe 800  
2581 TGATAGTCCACATTGAGTCTTTTGGCAACCGACTACTGAAAATTCACAGTCTACAAA 2640  
Db rValGluSerThrPheSerLeuPheGlyLysProThrThrGluAsnSerGlnSerThrLys 820  
2641 AGTGGAGGAAGCTTTAATCTTTACTACCAAGGAGGAGCAACAAAGACAGTAATGTGACA 2700  
Db sValGluGluAspPheAsnLeuThrThrLysGluGlyAlaThrLysThrValThrGlyGl 840  
2701 ACAGAACCTGTATTTGGCATTTTGAACGACTCCACAGATCAACAAATATAGATGCC 2760  
Db nGlnGluArgAspIleGlyIleGluArgAlaProGlnAspGlnThrAsnLysMetPr 860  
2761 CACATCAGAAATTAGGAAGAAAAGAGATACAAAATCAACTTCAGATTCTGAGATTATCTC 2820  
Db oThrSerGluLeuGlyArgLysGluAspThrLysSerThrSerAspSerGluIleIleSe 880  
2821 TGTGAGTGATACACAGAAATATGAGTGTGTACTGAGGCTACATATCAAAAGAAATAAA 2880  
Db rValSerAspThrGlnAsnTyrGluCysLeuProGluAlaThrThrGlnLysGluIleLys 900  
2881 GACAAACAATGCAAAATAGAGAGTCTCTGAAAAGCCTTCTCACTTTGAGCCTGCCAC 2940  
Db sThrThrAsnGlyLysIleGluSerProGluLysProSerHisPheGluProAlaTh 920  
2941 TGAATTCGAAAACCTCTGTTTCCAAATAAAGCCTTAGAATGGAAGAAATAACAAACATTGAG 3000

920 rGluMetGlnAsnSerValProAsnLysGlyLeuGluThrLysAsnLysGlnThrLeuAr 940  
3001 AGCAGATTCAACTACCTATCAAAAATCTTGATGACACTTCTCTTGTGAAAGAGGAAG 3060  
Db gAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeuProSerCysGluArgLysAr 960  
3061 CGAACTTTAAAAAGATAACTGTGAACAAAATTTACAGCAAAAATCGAACAAATGAAATAA 3120  
Db gGluLeuLysLysAspAsnCysGluGlnIleThrAlaLysMetGluGlnMetLysAsnLys 980  
3121 GTTTTGTGTACTACAAAAGAACTGTCAAGAGGAAAGAAATAAATACAGTGTAGAGAA 3180  
Db sPheCysValLeuGlnLysGluLeuSerGluAlaLysGluIleLysSerGlnLeuGluAs 1000  
3181 CCAAAAAGCTTAAATGGAACACAGACTCTGCACTGTGAGATTGCCCTTTAAATCAAGACA 3240  
Db nGlnLysAlaLysTrpGluGlnLeuCysSerValArgLeuThrLeuAsnGlnGluGl 1020  
3241 AGAGAGAGAGAAATGTGATATATTAAGAAAAGAAAATTAGACCCGAAGAGCAACTTAG 3300  
Db uGluLysArgArgAsnValAspIleLeuLysGluLysIleArgProGluGluGlnLeuAr 1040  
3301 GAAAAAGTTAGAGTGAACACCAACTTGAACAGACTCTCAGAAATACAAAGATATAGAAAT 3360  
Db gLysLysLeuGluValLysGlnLeuGluGlnThrLeuArgIleGlnAspIleGluLe 1060  
3361 GAAAACTGTAAACAAGTAATTTGAATCAGGTTTCTCACTCATGAAAGTGAATAATGATCT 3420  
Db uLysSerValThrSerAsnLeuAsnGlnValSerHisThrHisGluSerGluAsnAspLe 1080  
3421 CTTTCATGAAAATTTGATGTTTGAAGAGAAATTAATACTTTGAGGACATTAAGATTTTACA 3480  
Db uPheHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluValAlaTh 1100  
3481 ACTGAAACATCAACACAGCTGAAGGAAATTAATACTTTGAGGACATTAAGATTTTACA 3540  
Db rLeuLysHisGlnHisGlnValLysGluAsnLysTyrPheGluAspIleLysIleLeuGl 1120  
3541 AGAAAAGATGCTGAACCTTCAAAATGACCTTAAACTGAAACAGAAAACAGTACAAAAG 3600  
Db nGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGlnLysThrValThrLysAr 1140  
3601 GGCACTCTCAGTATAGAGAGCAGCTTAAAGTTTCTGAGCGAGAGAAACAGATGCTGACTTC 3660  
Db gAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAlaGluAsnThrMetLeuThrSe 1160  
3661 TAAATTGAAGAA 3673  
Db rLysLeuLysGlu 1164  
RESULT 2  
AAB84703  
ID AAB84703 standard; protein; 1011 AA.  
XX AC AAB84703;  
XX AC AAB84703;  
XX 17-SEP-2001 (first entry)  
DT XX  
DE XX Amino acid sequence of a human cancer associated antigen.  
XX XX Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.  
XX KW Homo sapiens.  
XX OS  
XX WO200147959-A2.  
XX PD 05-JUL-2001.  
XX 29-NOV-2000; 2000WO-US042334.  
XX PF  
XX 30-NOV-1999; 99US-00451739.  
XX PR 24-OCT-2000; 2000US-00602362.  
PR

XX (LUDW-) LUDWIG INST CANCER RES.  
 PA (SLOK) SLOAN KETTERING INST CANCER RES.  
 PA (CORR) CORNELL RES FOUND INC.  
 XX  
 XX Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;  
 XX WPI; 2001-441706/47.  
 XX  
 XX Isolated cancer associated nucleic acid molecule identified by SEREX  
 XX (serological identification of antigens by recombinant expression  
 XX cloning) technique, useful in nucleic acid based therapies to treat  
 XX cancer.  
 XX  
 XX Claim 83; Page 59-62; 62pp; English.  
 XX  
 CC The present sequence represents a human cancer associated antigen. The  
 CC sequence was identified using probes derived from the INGI gene. The INGI  
 CC gene is a tumour suppressor candidate gene. The cancer associated antigen  
 CC polynucleotides and polypeptides are useful for screening for the  
 CC possible presence of a pathological condition in a subject such as  
 CC cancer. The cancer associated antigen polypeptides are useful for  
 CC producing vaccines  
 CC  
 XX SQ Sequence 1011 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 1011  
 Score: 5173.00 Matches: 1011  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 80.23% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-602-362E-26 (1-3673) x AAB84703 (1-1011)  
 QY 641 ATGGTGGGAACACTGCTGCTCTATGGTCAGTCATCGAGGTGCAAAACAGGCTAGCCTC 700  
 Db 1 MetValAlaThrLeuLeuSerTyrGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20  
 QY 701 ACACCCCTTTTACTGGCCATACAGAAAGCAAGCAAACTGTGGAATTTTACTAACA 760  
 Db 21 ThrProLeuLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr 40  
 QY 761 AAAATGCAATGCAAAACGCTTTAATGAGTCTAAATGACAGCCCTCATGCTTGCCATA 820  
 Db 41 LysAsnAlaAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle 60  
 QY 821 TGTGAGGCTCATCAGATAGTCCGATGCTTCTTCAGCAAAATGTTGAGTCTTTGCT 880  
 Db 61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAla 80  
 QY 881 GAAGACATACATGGAATTAACCTCAGAACGTTATGCTGCTCGTGGAGTTAATTACATT 940  
 Db 81 GluAspIleHisGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100  
 QY 941 CATCAACAACCTTTGGACATATAGAAATTTACTTAAATCTCAAAATACCAATCCA 1000  
 Db 101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120  
 QY 1001 GAAGAACATCTACAGGAACACCTGATGAGGTGCACCCCTGGCGGAAAGAACACCTGAC 1060  
 Db 121 GluGlyThrSerThrGlyThrProAspGluAlaIleProLeuAlaGluArgThrProAsp 140  
 QY 1061 ACGGCTGAAGCTTGTGAAAAAACAACGTGACGAGGCTGCACGCTTGGTGGAGGAACG 1120  
 Db 141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160  
 QY 1121 TCTGCCAAATTAATGCTGGGGAAGCAACATCTGGAAGTTTGAACAGTCAACAGAA 1180  
 Db 161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180  
 QY 1181 GAAACACCTAGGAAATTTTGGAGCCTCAAAAGAAACATCTGAGAAATTTTCATGGCCA 1240

Db 181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerTyrPro 200  
 QY 1241 GCAAAGAAAGACTAGGAAGATCACATGGAGGAGAAAGAAACATCTGTAAAGACTGAA 1300  
 Db 201 AlalysGluArgSerArgLysIleThrTyrGluGluLysGluThrSerValLysThrGlu 220  
 QY 1301 TGGTGCAGAGTAACACCTAATAAAGTCAAGTTTGGAAAAGGACATCTATATG 1360  
 Db 221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240  
 QY 1361 ATTGCATGCTCTACAAAAGAAACATCTACAAAAGCAAGTACAAATGTGATGTGACTTCT 1420  
 Db 241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSerSer 260  
 QY 1421 GTAGAGCTATATTCAGTCTTTTGGCACAGGACTATTGAAAATTCACAGTGTCAAAA 1480  
 Db 261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280  
 QY 1481 GTTGAGGAAGACTTTAATCTTGTACCAAGATTATCTTAAGAGTGTCTGCACAGAATTAT 1540  
 Db 281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300  
 QY 1541 ACGTGTTTACCTGATGCTACATATCAAAAAGATATCAAAACAAATAAATCACAAAATAGAA 1600  
 Db 301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320  
 QY 1601 GATCAGATGTTCCCATCAGAAATCCAAACGAGAGGAAGATGAAGAAATATCTTGGGATCT 1660  
 Db 321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluLysSerTyrSerTyrSer 340  
 QY 1661 GGGAGTCTCTTTCAGAGTCTTGCAAAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720  
 Db 341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360  
 QY 1721 AAGTATGAGATAAATAGAGAGTAGAGAGTCTTCTGAGAGGCATCTCCCTCAAG 1780  
 Db 361 LysValMetGluIleAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380  
 QY 1781 CTGTGGGATTCTCAGAGTCCCTGTGAGCGGTTTCAAGAAGGATGTGATTTACCCAAA 1840  
 Db 381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400  
 QY 1841 ACATTGAGACAGCTCAGATGTTCCCATCAGAAATCCAAACAAAGAGACGATGAAGAAAT 1900  
 Db 401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspGluGluAsn 420  
 QY 1901 TCTTGGGATTCTCAGAGTCCCTGTGAGCGGTTTCAAGAAGGATGTGATTTACCCAAA 1960  
 Db 421 SerTyrAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440  
 QY 1961 GCTACACATCAAAAAGAAATTCGATACCTTAAGTGGAAATTTAGAGAGTCTCCTGTGATAA 2020  
 Db 441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluGluSerProValLys 460  
 QY 2021 GATGCTCTCTGAGGCTACCTGTGGAAGGAAGTTTCTTCCAAATAAAGCCTTAGAA 2080  
 Db 461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480  
 QY 2081 TTAAGGACAGAGAAACATTCAAAAGCAGAGTCTCTGTGATAAAGATGCTTCTTGAAGCCT 2140  
 Db 481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500  
 QY 2141 ACCTGTGGAGGAAGGTTTCTCTTCCAAATAAGCCTTAGAATTAAGGACAGAGAAACA 2200  
 Db 501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520  
 QY 2201 CTCAAGACAGAGTCTCTCTGATATGATGCTTCTTGAAGCCTTACCTGTGGAAGGAAGT 2260  
 Db 521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540  
 QY 2261 TCTCTTCCAAATAAGCTTTAGAAATGAAGGACAGAGAAACATTCAAAGCAGCTCAGATG 2320

Db 541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaAlaGlnMet 560  
 QY 2321 TTCCATCAGAAATCCAAACAAAGAGATGATGAAGAAAATTTCTGGGATTTTGGAGTTTC 2380  
 Db 561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580  
 QY 2381 CTTGAGACTCTCTACAGAATGATGTGTGTTTACCAGGCTACACATCAAAAAGAAATTC 2440  
 Db 581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600  
 QY 2441 GATACCTTAAGTGAAGAAATTAAGAGAGTCTCCTGATAAGATGGTCTCTGAAAGCTTACC 2500  
 Db 601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620  
 QY 2501 TGTGGAATGAAAATTTCTCTCCAAATAAGCCCTTAGAATTGAAGGACAGAGAAACATTC 2560  
 Db 621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640  
 QY 2561 AAAGACAGAGGTGTCAGTCTGTAGAGTCCACATTCAGTCTCTTTGGCAACCCAGCTACT 2620  
 Db 641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660  
 QY 2621 GAAATTCACAGTCTACAAAAGTTGAGGAAGACTTTAATCTTACTTACCAGAGGGAGCA 2680  
 Db 661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680  
 QY 2681 ACAAGACAGTAACCTGGACACAGAGACGTGATATTTGGCATTTATGACGAGCTCCACAA 2740  
 Db 681 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleGluArgAlaProGln 700  
 QY 2741 GATCAACAAATAAGATGCCACATCAGAAATTAGAAGAAAAGAGATACAAAATCAACT 2800  
 Db 701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720  
 QY 2801 TCAGATCTGAGATTATCTGTGATGATACACAGAAATATGAGTGTTTACCTGAGGCT 2860  
 Db 721 SerAspSerGluIleSerValSerAspThrGlnAsnTyrgluCysLeuProGluAla 740  
 QY 2861 ACATATCAAAAGAAATAAAGACAAACAAATGAGAGAGTCTCCTGAAAAGCCT 2920  
 Db 741 ThrTyrglnLysGluLeuLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760  
 QY 2921 TCTACTTTGAGCCTGCCACTGAAATGCAAAACTCTGTCCAAATPAAAGGCTTGAATGG 2980  
 Db 761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780  
 QY 2981 AGAATAACAACATTTGAGACAGATTCACCTACCTATCAAAATCTTGATGTCACCTT 3040  
 Db 781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800  
 QY 3041 CTTCTTTGTGAAGAGGAGGAACTTAAAAAGATAACTGTGAACAAATACAGCAAAA 3100  
 Db 801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820  
 QY 3101 ATGGAACAAATGAAAATTAAGTTTGTGTACTACAAAAGGAAGTCTCAGAGCGGAAGAA 3160  
 Db 821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840  
 QY 3161 ATAAATACAGTTAGAGAACCAAAAAGCTAAATGGAAACAGAGCTCTGAGTGTGAGA 3220  
 Db 841 IleLysSerGlnLeuGluAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860  
 QY 3221 TTGCCTTTAAATCAAGAAGAGAGAGAGAAATGTGATATTTAAAGAAAAATTT 3280  
 Db 861 LeuProLeuAsnGlnGluGluLysArgArgAsnValAspIleLeuLysGluLysIle 880  
 QY 3281 AGACCCGAGAGCACTTAGGAAAAGTTAGAGTGAACACCACTTGAACAGACTCTC 3340  
 Db 881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGlnThrLeu 900  
 QY 3341 AGAATAACAAGATATAAGATTTGAAAAGTGTAAACAGTAAATTTGAATCAGGTTTCTCACAT 3400  
 Db 901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920

QY 3401 CATGAAAGTGAATAATGATCTCTTCATGAAATTCATGATCTTGAATAAGGAAATTCGCATG 3460  
 Db 921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940  
 QY 3461 CTAAAACTGCAAGTACCCACACTGAAACATCAACACAGGTGAAGGAAATATAATCTTT 3520  
 Db 941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrrPhe 960  
 QY 3521 GAGGACATTAAGATTTTACAAGAAAAGATGCTGAACCTCAAAATGACCTTAAACTGAAA 3580  
 Db 961 GluAspIleLysIleLeuGlnGluLysAsnAlaGluLeuGlnMetThrLeuLysLys 980  
 QY 3581 CAGAAACAGTAAACAAAAGGGCATCTCAGTATAGAGACAGCTTAAAGTCTTGACGGCA 3640  
 Db 981 GlnLysThrValThrLysArgAlaSerGlnTyArgGluGlnLeuLysValLeuThrAla 1000  
 QY 3641 GAGAACACAGATGCTGACTTCTTAAATGAAGGAA 3673  
 Db 1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011  
 RESULT 3  
 ABU37789  
 ID ABU37789 standard; protein; 1239 AA.  
 XX AC ABU37789;  
 XX 15-MAY-2003 (first entry)  
 DT Human tumour-related protein - SEQ ID No 577.  
 XX Human; vaccine; gene therapy; T cell stimulation; T cell expansion;  
 KW tumour; breast cancer; cancer; immune response stimulation.  
 XX Homo sapiens.  
 OS W0200283956-A1.  
 PN 24-OCT-2002.  
 XX 15-APR-2002; 2002WO-US012378.  
 PF 13-APR-2001; 2001US-00834759.  
 PR 07-DEC-2001; 2001US-00007805.  
 PR 13-FEB-2002; 2002US-00076622.  
 XX (CORI-) CORIXA CORP.  
 PA Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;  
 PI Mitcham JU, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;  
 PI Vedvick TS, McNeill PD, Durham M;  
 XX WPI; 2003-103376/09.  
 XX New polypeptide and polynucleotide useful for stimulating and/or  
 PT expanding T cells specific for a tumor protein and treating breast  
 PT cancer.  
 XX Example 13; Page 359-363; 375pp; English.  
 PS The invention comprises a method of stimulating and/or expanding T cells  
 CC specific for a tumour protein. The invention further comprises human  
 CC nucleic acids and proteins that are associated with tumours (e.g. breast  
 CC cancer). The method and sequences of the invention are useful for  
 CC stimulating and/or expanding T cells specific for a tumour protein,  
 CC detecting the presence of cancer, stimulating an immune response in a  
 CC patient and treating breast cancer. The present amino acid sequence  
 CC represents a human tumour-related protein  
 XX Sequence 1239 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 1239

Score: 5173.00 Matches: 1011  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 80.23% Indels: 0  
 DB: 6 Gaps: 0

US-09-602-362E-26 (1-3673) x ABU37789 (1-1239)

QY 641 ATGGTGGCAACACTGCTGCTCTATGTCAGTCATGAGGTGCAGAAACAGGCTAGCCCTC 700  
 DB 1 MetValAlaThrLeuLeuSerGlyAlaValIleGluValGlnAsnLysAlaSerLeu 20  
 QY 701 ACACCCCTTTTACTGGCCATACAGAAAGAACAGCAAACTGTGGAATTTTACTAACA 760  
 DB 21 ThrProLeuLeuAlaIleGlnLysArgSerLysGlnThrValGluPheLeuLeuThr 40  
 QY 761 AAAATGCAATGCAACCACTTAAATGAGTCTAAATGACAGCCCTCATGCTGCCATA 820  
 DB 41 LysAsnAlaAsnAlaPheAsnGluSerLysCysThrAlaLeuMetLeuAlaIle 60  
 QY 821 TGTGAAGGCTCATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGCTTTTGT 880  
 DB 61 CysGluGlySerSerGluIleValGlyMetLeuLeuGlnAsnValAspValPheAla 80  
 QY 881 GAAGACATACATGGAATTAATCTGAGAACGTTATGCTGCTGCTGGAGTTAATTACATT 940  
 DB 81 GluAspIleGlyIleThrAlaGluArgTyrAlaAlaAlaArgGlyValAsnTyrIle 100  
 QY 941 CATCAACAACTTTTGGAAACATATACAAATTTACCTAAAAATCCCTCAAAATACCAATCCA 1000  
 DB 101 HisGlnGlnLeuLeuGluHisIleArgLysLeuProLysAsnProGlnAsnThrAsnPro 120  
 QY 1001 GAAGAACATCTACAGAAACACTGATGAGGTGCACCTGCGGGAAGAACACCTGCAC 1060  
 DB 121 GluGlyThrSerThrGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAsp 140  
 QY 1061 ACGGCTGGAAGCTTGTGGAAGAACACCTGACAGGCTGCACGCTTGGTGGAGGACG 1120  
 DB 141 ThrAlaGluSerLeuLeuGluLysThrProAspGluAlaAlaArgLeuValGluGlyThr 160  
 QY 1121 TCTGCCAAATTCATGCTGGGGAAGCAACATCTGGAAGTTTGAACAGTCAACAGAA 1180  
 DB 161 SerAlaLysIleGlnCysLeuGlyLysAlaThrSerGlyLysPheGluGlnSerThrGlu 180  
 QY 1181 GAACACCTAGAAATTTTGGGCTACAAAGAACATCTGAGAAATTTTCAATGCGCA 1240  
 DB 181 GluThrProArgLysIleLeuArgProThrLysGluThrSerGluLysPheSerThrPro 200  
 QY 1241 GCAAAAGAAAGATCTAGGAAGATCATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1300  
 DB 201 AlaLysGluArgSerArgLysIleThrTrpGluGluLysGluThrSerValLysThrGlu 220  
 QY 1301 TCGTGGCGAGGAGTAAACCTTAATTAAGTAACTGAGTTTGGGAAAGGAGCACTAATATG 1360  
 DB 221 CysValAlaGlyValThrProAsnLysThrGluValLeuGluLysGlyThrSerAsnMet 240  
 QY 1361 ATTGCATGCTCTACAAAGAACATCTCAAAAGCAAGTACAAATGTGATGTGAGTTCT 1420  
 DB 241 IleAlaCysProThrLysGluThrSerThrLysAlaSerThrAsnValAspValSer 260  
 QY 1421 GTAGAGCTTAATTCAGTCTTTTGGCACACGACTATTGAAATTCACAGTGTACAAA 1480  
 DB 261 ValGluProIlePheSerLeuPheGlyThrArgThrIleGluAsnSerGlnCysThrLys 280  
 QY 1481 GTTGAAGGAGCTTAACTTGTCTACCAAGATTAATCTTAAGTGTGCTGCACAGATTAT 1540  
 DB 281 ValGluGluAspPheAsnLeuAlaThrLysIleIleSerLysSerAlaAlaGlnAsnTyr 300  
 QY 1541 ACGTGTTCCTGATGCTACATATCAAAAGATATCAAAACAAATAAATCAAAATAGAA 1600  
 DB 301 ThrCysLeuProAspAlaThrTyrGlnLysAspIleLysThrIleAsnHisLysIleGlu 320  
 QY 1601 GATCAGATGTTCCCATCAGAAATCCAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1660

DB 321 AspGlnMetPheProSerGluSerLysArgGluGluAspGluGluThrSerTrpAspSer 340  
 QY 1661 GGGAGTCTCTTTGAGATCTTCGCAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAG 1720  
 DB 341 GlySerLeuPheGluSerSerAlaLysThrGlnValCysIleProGluSerMetTyrGln 360  
 QY 1721 AAAGTAATGAGATATAATAGAGAGTAGAGAGCTTCCTCAGAGCCATCTGCTCTCAAG 1780  
 DB 361 LysValMetGluLeuAsnArgGluValGluGluLeuProGluLysProSerAlaPheLys 380  
 QY 1781 CTGCTCCCTNGAAATGCAAAAGACTGTTCCAAATAAAGCCCTTTCGAATGAAGATGAACA 1840  
 DB 381 ProAlaValGluMetGlnLysThrValProAsnLysAlaPheGluLeuLysAsnGluGln 400  
 QY 1841 ACATTGAGACAGCTCAGATGTTCCCATCAGATCCCAACAAAGAGCAGATGAAGAAAT 1900  
 DB 401 ThrLeuArgAlaAlaGlnMetPheProSerGluSerLysGlnLysAspGluGluAsn 420  
 QY 1901 TCTTGGGATTTCTGAGAGTCCCTCTGAGACGTTTTCACAGAGGATGTTATTTACCCAAA 1960  
 DB 421 SerTrpAspSerGluSerProCysGluThrValSerGlnLysAspValTyrLeuProLys 440  
 QY 1961 GCTACACATCAAAAGAAATTCGATACCTTAAGTGGAAATTTAGAGAGTCTCTGTTAA 2020  
 DB 441 AlaThrHisGlnLysGluPheAspThrLeuSerGlyLysLeuGluSerProValLys 460  
 QY 2021 GATGGTCTTCTGAGGCTTACCTGCGAAGGAAAGTTTCTTCCAAATAAAGCCTTAGAA 2080  
 DB 461 AspGlyLeuLeuLysProThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGlu 480  
 QY 2081 TTAAGACACAGAAACATTCAAAGCAGAGTCTCTGATAAAGATGGTCTTCTGAAGCCT 2140  
 DB 481 LeuLysAspArgGluThrPheLysAlaGluSerProAspLysAspGlyLeuLeuLysPro 500  
 QY 2141 ACCTGTGGAAGGAAAGTTTCTTCCAAATAAAGCCTTAGAATTAAGGACACAGAGAAACA 2200  
 DB 501 ThrCysGlyArgLysValSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThr 520  
 QY 2201 CTCAAGACAGAGTCTCTGATATGATGCTCTTCTGAAGCCTACCTGTGGAAGGAAAGTT 2260  
 DB 521 LeuLysAlaGluSerProAspAsnAspGlyLeuLeuLysProThrCysGlyArgLysVal 540  
 QY 2261 TCTTCTCCAAATAAAGCCTTTCGAAATTAAGGACAGAGAAACATTCAGACAGTCTCAGTG 2320  
 DB 541 SerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPheLysAlaAlaGlnMet 560  
 QY 2321 TTCCCATCAGAAATCCAAACAAAGGATGATGAAGAAATTTCTTGGGATTTTCAGAGTTTC 2380  
 DB 561 PheProSerGluSerLysGlnLysAspAspGluGluAsnSerTrpAspPheGluSerPhe 580  
 QY 2381 CTGACAGCTCTTACAGAAATGATGTGCTTTTCCCAAGGCTACACATCAAAAGAAATTC 2440  
 DB 581 LeuGluThrLeuLeuGlnAsnAspValCysLeuProLysAlaThrHisGlnLysGluPhe 600  
 QY 2441 GATACCTTAAGTGGAAATTAAGAGAGTCTCTGATTAAGATGCTTCTCTGAAGCTTACC 2500  
 DB 601 AspThrLeuSerGlyLysLeuGluGluSerProAspLysAspGlyLeuLeuLysProThr 620  
 QY 2501 TGTGGAATGAAATTTCTTCTTCCAAATAAAGCCTTAGAATTCGAAGACAGAGAAACATTC 2560  
 DB 621 CysGlyMetLysIleSerLeuProAsnLysAlaLeuGluLeuLysAspArgGluThrPhe 640  
 QY 2561 AAAGACAGAGGATGTGAGTCTCTGATGCTCCATTCAGTTCAGTCTTTTGGCAAAACCGACTACT 2620  
 DB 641 LysAlaGluAspValSerSerValGluSerThrPheSerLeuPheGlyLysProThrThr 660  
 QY 2621 GAAATTCACAGTCTACAAAGTTCAGGAAGACTTTAATCTTACTACCAAGGAGGAGCA 2680  
 DB 661 GluAsnSerGlnSerThrLysValGluGluAspPheAsnLeuThrThrLysGluGlyAla 680  
 QY 2681 ACAAGACAGTAGTAACGACCAACAGCAACGTCGATATTTGGCATTTATGAACGAGCTCCAAA 2740

Db 691 ThrLysThrValThrGlyGlnGlnGluArgAspIleGlyIleIleGluArgAlaProGln 700  
Qy 2741 GATCAAAATAAGATGCCACATCAGATTAAGAGAGAAAGATACAAAATCAACT 2800  
Db 701 AspGlnThrAsnLysMetProThrSerGluLeuGlyArgLysGluAspThrLysSerThr 720  
Qy 2801 TCAGATCTGAGATTACTCTGTCAGTATACACAGAAATATGAGTGTTCACCTGAGGCT 2860  
Db 721 SerAspSerGluIleIleSerValSerAspThrGlnAsnTyrGluCysLeuProGluAla 740  
Qy 2861 ACATATCAAAAGAAATAAAGACAAACAAATAGCAAGAGTCTCTGAAAGCCT 2920  
Db 741 ThrTyrGlnLysGluIleLysThrThrAsnGlyLysIleGluGluSerProGluLysPro 760  
Qy 2921 TCTACATTTAGCTGCCACTGAATGCCAAATCTCTGTCACATTAAGGCTTAGATGG 2980  
Db 761 SerHisPheGluProAlaThrGluMetGlnAsnSerValProAsnLysGlyLeuGluTrp 780  
Qy 2981 AAGAATAACAAACATTCAGAGCAGATTCACACTACCTCTATCAAAATCTTGATGCACTT 3040  
Db 781 LysAsnLysGlnThrLeuArgAlaAspSerThrThrLeuSerLysIleLeuAspAlaLeu 800  
Qy 3041 CCTCTTCTGAAAGAGGAGGAATTAATAAGATACTGTGAACAAATTAACAGCAAAA 3100  
Db 801 ProSerCysGluArgGlyArgGluLeuLysLysAspAsnCysGluGlnIleThrAlaLys 820  
Qy 3101 ATGGAAACAAATCAAAATAAGTTTGTCTACTACAAAAGCACTGTGAGAAAGCAAGAA 3160  
Db 821 MetGluGlnMetLysAsnLysPheCysValLeuGlnLysGluLeuSerGluAlaLysGlu 840  
Qy 3161 ATAAATCACAGTTAGAGAACAAAGAACTAAATGGGAACAGAGCTCTGAGTGTGAGA 3220  
Db 841 IleLysSerGlnLeuGlnAsnGlnLysAlaLysTrpGluGlnGluLeuCysSerValArg 860  
Qy 3221 TTGCTTTTAATCAGAGACAGAGAGAGAAATGTCGATATATTAAGAGAAAAATT 3280  
Db 861 LeuProLeuAsnGlnGluGluLysArgGlnValAspIleLeuLysGluLysIle 880  
Qy 3281 AGACCCGAGAGCAACTTAGGAAAAAGTTAGAAGTGAACACCACTTTGAACAGACTCTC 3340  
Db 881 ArgProGluGluGlnLeuArgLysLysLeuGluValLysHisGlnLeuGlnThrLeu 900  
Qy 3341 AGAATACAGATATAGATTGAAGTGTACCAAGTAAATTTGAATCAGGTTTCTCACACT 3400  
Db 901 ArgIleGlnAspIleGluLeuLysSerValThrSerAsnLeuAsnGlnValSerHisThr 920  
Qy 3401 CATGAAAGTGAATGATCTCTTTCATGAAATTCATGTTGAAAGAGCAAAATGCCCATG 3460  
Db 921 HisGluSerGluAsnAspLeuPheHisGluAsnCysMetLeuLysLysGluIleAlaMet 940  
Qy 3461 CTAATACTGGAAGTAGCCACACTGAACATCAACACCCAGGTGAAGGAAAAATAACTTTT 3520  
Db 941 LeuLysLeuGluValAlaThrLeuLysHisGlnHisGlnValLysGluAsnLysTyrPhe 960  
Qy 3521 GAGGACATTAGATTTTACAGAAAGAGATGCTGAACCTCAATGACCTTAAACTGAAA 3580  
Db 961 GluAspIleLysIleLeuGlnGlnLysAsnAlaGluLeuGlnMetThrLeuLysLysLys 980  
Qy 3581 CAGAAAAACAGTAACAAAAAGGCATCTCAGTATAGAGACGACTTAAAGTTCTGACGCCA 3640  
Db 981 GlnLysThrValThrLysArgAlaSerGlnTyrArgGluGlnLeuLysValLeuThrAla 1000  
Qy 3641 GAGAACACATGCTGACTTCTTAATTAAGGAA 3673  
Db 1001 GluAsnThrMetLeuThrSerLysLeuLysGlu 1011

RESULT 4  
ID AAB84702 standard; protein; 1341 AA.  
AC AAB84702;  
XX  
XX AAB84702;  
XX  
DT 17-SEP-2001 (first entry)

XX DE Amino acid sequence of a human cancer associated antigen.  
XX KW Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.  
XX OS Homo sapiens.  
XX PN WC200147959-A2.  
XX PD 05-JUL-2001.  
XX 29-NOV-2000; 2000WO-US042334.  
XX 30-NOV-1999; 99US-00451739.  
XX 24-OCT-2000; 2000US-00602362.  
XX (LUDM-) LUDWIG INST CANCER RES.  
XX (SLOK) SLOAN KETTERING INST CANCER RES.  
XX (CORR) CORNELL RES FOUND INC.  
XX Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;  
XX WPI; 2001-441706/47.  
XX Isolated cancer associated nucleic acid molecule identified by SEREX  
XX (serological identification of antigens by recombinant expression  
XX cloning) technique, useful in nucleic acid based therapies to treat  
XX cancer.  
XX Claim 83; Page 53-57; 62pp; English.  
XX The present sequence represents a human cancer associated antigen. The  
XX sequence was identified using probes derived from the INGI gene. The INGI  
XX gene is a tumour suppressor candidate gene. The cancer associated antigen  
XX polynucleotides and polypeptides are useful for screening for the  
XX possible presence of a pathological condition in a subject such as  
XX cancer. The cancer associated antigen polypeptides are useful for  
XX producing vaccines  
XX SQ Sequence 1341 AA;  
Alignment Scores:  
Pred. No.: 2,38e-273 Length: 1341  
Score: 3458.00 Matches: 740  
Percent Similarity: 65.58% Conservative: 83  
Best Local Similarity: 58.96% Mismatches: 159  
Query Match: 53.63% Indels: 273  
DB: 4 Gaps: 12  
US-09-602-362E-26 (1-3673) x AAB84702 (1-1341)  
Qy 349 ATGACACTAGGAGAGAGCCGTCACCTGAACAAAGAGATATGAAGAGAGACTGCT 408  
Db 1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20  
Qy 409 CTACACTGGGCGCTGTGTCAATGGCCATGCCANAAAGTAGTAACATTCTGTGTAGACAGAAAG 468  
Db 21 LeuHisTrpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40  
Qy 469 TGCNGCTTAATGTCTTGTATGGCGAAGGAGGACCTCTGTATGAGGCTTACATGTC 528  
Db 41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60  
Qy 529 GAGAGGGAAGCTTTGTGCAATATTCATAGATGCTGTGCTGATCTAAATATATGTAGA 588  
Db 61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80  
Qy 589 TGTGTATGGCAACACGGCTCTCCATTATGCCGTTTATGTAGAGAAATTTATATGTGTGC 648  
Db 80 pValTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100  
Qy 649 AACACTGCTGCTTATGTCAGTCATCGAGGTGCAAAACAGGCTAGCCTCACACCCCT 708

Db 100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120  
Qy 709 TTACTGGCCATACAGAAAGAGCAAGCAACTGTGGATTTTACTTAACAAAATGCG 768  
Db 120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140  
Qy 769 AAATGCAAAACGCAATTTAATGAGTCTAAATGCGACAGCCCTCATGCTTGCCTATGTGAAGG 828  
Db 140 aAsnAlaAsnAlaValAsnLysThrLysCysThrAlaLeuMetLeuAlaValCysHisG 160  
Qy 829 CTATCAGAGATAGTGGGATGCTTCTTTCAGCAAAATGTGACGCTTTGGCTGAAGACAT 888  
Db 160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspI 180  
Qy 889 ACATGAATAACTGCAGAACGTTATGCTGCTGCTGGAGTTAATTACATTCATCAACA 948  
Db 180 eCysGlyValThrAlaGluHisTyrrAlaValThrCysGlyPheHisHisIleGluG 200  
Qy 949 ACTTTGGAAACATATACGAAATTTACCTAAAATCTCAAAATACCAATCCAGAGGAAC 1008  
Db 200 nIleMetGluTyrrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGly 220  
Qy 1009 ATCTACAGAAACACCTGATGAGGCTGCACCCCTGGCG 1045  
Db 220 rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaG 240  
Qy 1046 1062  
Db 240 uSerLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspTh 260  
Qy 1063 GGCTGAACTGCTGCGAAACACCTGACAGAGCTGCACGCTTGGTGAGGGAACGTC 1122  
Db 260 rAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSe 280  
Qy 1123 TGCCTAAATCAATGTCTCGGAAAGCAACATCTCGAAAGTTTGAACAGTCAACAGAAGA 1182  
Db 280 rAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluG 300  
Qy 1183 ACACCTAGGAAATTTGAGCGCTACAAAGAAACATCTGAGAAATTTGATGCGCA 1240  
Db 300 uThrProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAl 320  
Qy 1240 1240  
Db 320 aLysGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMe 340  
Qy 1241 1260  
Db 340 tSerProAlaLysGluThrSerGluLysPheThrTrpAlaLysGlyArgProArgLy 360  
Qy 1261 GATCATGGGAGGAAAGAAACATCTCTAAGACTGAATCCGTCGCGAGGAGTAACACC 1320  
Db 360 sIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSe 380  
Qy 1321 TAATAAACTGAAGTTTGGAAAAAGAACATCTAATATGATGTGATGCTCTACAAAAGA 1380  
Db 380 rAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysG 400  
Qy 1381 ACATCTACAAAGCAAGTACAAATGTGGATGTGAGTCTGTAGAGCCTATATTCAGTCT 1440  
Db 400 uSerSerThrLysAlaSerAlaAsn 408  
Qy 1441 TTTTGGCACCGACTATTGAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCT 1500  
Db 408 408  
Qy 1501 TGCTACCAAGATTATCTTAAGAGTGTGCGACAGAAATATACGTGTTTACTGATGCTAC 1560  
Db 408 408  
Qy 1561 ATATCAAAAAGATATCAAAAACAATAATCAAAAATAGAGATCAGATGTTCCCATCAGA 1620  
Db 409 409  
AspGlnArgPheProSerG 415

Qy 1621 ATCCARACGAGAGAGATGAAGATATTTCTTGGATTTCTGGGAGTCTCTTTGAGAGTTC 1680  
Db 415 uSerLysGlnGluAspGluGluTyrrSerCysAspSerArgSerLeuPheGluSerSe 435  
Qy 1681 TGCAAAAGACTCAAGTGTGTATACCTGAGTCTATGTATATCAGAAAAGTATGAGAGATAATAG 1740  
Db 435 rAlaLysIleGlnValCysIleProGluSerIleTyrrGlnLysValMetGluLeuAsnAr 455  
Qy 1741 AGAAGTAGAGAGCTTCTCAGAGAGCCATCTGCTTCAAGCTTCCGCTGCAAGTCAAAA 1800  
Db 455 gGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAs 475  
Qy 1801 GACTGTTTCCAAATAAAGCCTTGAATTTGAAGATGAACAAACATTGAGAGCAGCTCAGAT 1860  
Db 475 nSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProme 495  
Qy 1861 GTTCCCATCAGATCCAAACAAAGGACGATGAAGAAAATTTCTGGATTTCTGAGAGTCC 1920  
Db 495 tPheProProGluSerLysGlnLysAspTyrrGluGluAsnSerTrpAspSerGluSerLe 515  
Qy 1921 CTGTGACACGCTTTCAGAGAGGATGTATTATTACCCCAAGCTACACATCAAAAGAAAT 1980  
Db 515 uCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluI 535  
Qy 1981 CGATACCTTAAGTGAAGAAATTTAGAAGAGTCTCTGTTTAAAGATGCTTCTGAAGCCTAC 2040  
Db 535 eAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaTh 555  
Qy 2041 CTGTGAGAGAAATTTCTTCCCAATTAAGCCTTAGAATTAAGGACAGAGAAACATT 2100  
Db 555 rCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPh 575  
Qy 2101 CAAAGCAGAGTCTCTCAAGCTACCTGCTGGAAGAAAGTTTCTTCCCAATAAAGCTTT 2160  
Db 575 eLysAlaGluProProGlyLysProSerAlaPheGluPro 588  
Qy 2161 TCTTCCAAATAAGCCTTAGAATTAAGGACAGAGAAACACTCAAGAGAGAGTCTCCTGA 2220  
Db 589 594  
Qy 2221 TAATGATGGTCTTCTCAAGCCTACCTGTGGAAGAAAGTTTCTTCCCAATAAAGCTTT 2280  
Db 595 601  
Qy 2281 AGAATTAAGGACAGAGAAACATTAAGAGAGCTCAGATGTTCCCATCAGATCCAAACA 2340  
Db 601 uGluLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysG 621  
Qy 2341 AAAGGATGATGAAGAAATTTCTGGATTTTCTGAGAGTCTTCTTCTGAGACTCTCTACAGAA 2400  
Db 621 nLysAspTyrrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLy 641  
Qy 2401 TGATGTGTGTTTACCAAGCTACACATCAAAAGAAATTCGATACCTTAAGTGAAGAAAT 2460  
Db 641 sAspValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLe 661  
Qy 2461 AGAAGAGTCTCTGTATAAGATGCTTCTGAGAGCTTCTGAGAGCTTCTGAGAGTCTCT 2520  
Db 661 uGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerI 681  
Qy 2521 TCCAAATAAAGCCTTAGAATTTGAAGACAGAGAAACATTCAAAGCAGAGAGATGTGAGTTC 2580  
Db 681 eProThrLysAlaLeuLeuMetAspMetGlnThrPheLysAlaGluProProGluLy 701  
Qy 2581 TGTAGAGTCCATTCAGTCTTTTGGCAACCGACTACTGAAATTCACAGTCTACAAA 2640  
Db 701 sProSerAlaPhe 715  
Qy 2641 AGTTGAGGAAGACTTAAATCTTACTTACTCAAGGAGGAGCAACAAACAGTAACTGGACA 2700  
Db 716 723  
ProAsnLysAlaLeuGluLeuLy





Pred. No.:	2.38e-273	Length:	1341
Score:	3458.00	Matches:	740
Percent Similarity:	65.58%	Conservative:	83
Best Local Similarity:	58.96%	Mismatches:	159
Query Match:	53.63%	Indels:	273
DB:	5	Gaps:	12

US-09-602-362E-26 (1-3673) X ABJ05537 (1-1341)

349	ATGACAGT	AGGGAAGAGCCCGTCAACCTGAA	CAAAAGAGATGATGAAGAGAGCATGCT	408
QY				
1	MetThr	LeuArg	LeuGly	Thr:::leu:::asn:::ile:::gln:::asp:::ala:::gln:::arg:::thr:::ala:::20
QY	CTACACTGGCGCTGTGCTCAATGGCCATGCANAACTAGTAACTATTCCTGGT	TAGACAGAAAG	468	
Db				
21	LeuHis	TrpAla	CysVal	asnGlyHisGluGluValValThrPheLeuValaspArgLys40
QY	TGCNGCTTAAATGTCCTTGATGGGGAAGGAGGACACCTCTGTAGTGAAGGCTCTTACAATGC	528		
Db				
41	CysGln	LeuAsp	ValLeu	aspGlyGluHisArgThrProLeuMetLysalaLeuGlnCys60
QY	GAGAGGGAGCTTGTGCAAAATATTCATAGATCTGGTGTGATCTAAATATATCTAGA	588		
Db				
61	HisGln	GluAla	-CysAla	asnIleLeuileaspSerGlyalaaspIleasnLeuValas80
QY	TGTGATGGCAACACGGCTCTCCATTATGCGTTTATAGTGAGAAATTTATTAAATGGTGC	648		
Db				
80	pVal	TrpGly	AsnMet	AlaLeuHisIstyrAlaValItyrSerGluIleLeuSerValValAl100
QY	AACACTGCTGCTTATGCTGAGTCAGTCATCGAGGTGC	CAAAACAAGGCTAGCCTCACACCCCT	708	
Db				
100	ala	LeuLeuSer	HisGlyAlaValIleGluValHisAsn	LysAlaSerLeuThrProLe120
QY	TTTACTGGCCATPACAGAAAGCAAGCAACACTGTGGAAATTTTACTAACAAAAATGC	768		
Db				
120	uLeu	LeuSer	IleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl140	
QY	AAATCCAAACGATTTAATGAGTCTAAATGCACAGCCCTCATGCTTGC	CCATATGTGAGG	828	
Db				
140	aa	snAla	asnAlaVal	asnLysThrLysCysThrAlaLeuMetLeuAlaValCysHisGlu160
QY	CTCATCAGAGATPAGTCGGCATGCTTCTTCAGCAAAATGTTGACGCTTTTGC	TGCAAGACAT	888	
Db				
160	ySer	SerGlu	IleValGlyMetLeuLeuGlnAsnValaspVal	pheAlaalaaspil180
QY	ACATGGAATACTGCAGACAGCTTATGCTGCTGCTGGAGTTAATATCATTCATCAACA	948		
Db				
180	eCys	GlyValThrAlaGluHisIstyrAlaValThrCysGlyPheHisIleHisGluGlu	200	
QY	ACTTTTGGACATATACGAAATTAACCTTAAATATCTCTCAAAATACCAATCC	CAGAGGAAC	1008	
Db				
200	nileMet	GluTyrIleArgLysLeuSerLysAsnHisGln	asnThrAsnProGluGlyTh220	
QY	ATCTACAGGAACACTGTATGAGGTGCACCCCTGGCG	-----	1045	
Db				
220	rSer	AlaGlyThrProaspGluAlaAlaProLeuAlaGluArgThrProaspThrAlaGlu	240	
QY	-----	-----	-----	1046
Db				
240	uSer	LeuValGluLysThrProaspGluAlaAlaProLeuValGluArgThrProaspTh	260	
QY	GCCTGAAAGCTTGTGGAAAAACACCTGCAGAGGCTGCACGCTTGCTGGAGGGAA	CGTC	1123	
Db				
260	rAla	GluSerLeuValGluLysThrProaspGluAlaAlaSerLeuValGluGlyThrSe	280	
QY	TGCCAAAAATTCATGCTCTGGGGAAGCAACATCTGGAAGTTTGAACAGTCAAC	GAGAAGA	1182	
Db				
280	rAsp	LysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGlu	300	
QY	AAACACTAGGAAATTTTGAGCGCTACAAAGAAACATCTGAGAAATTTTCATGCCA	-----	1240	
Db				
300	uThr	ProargGluIleThrSerProAlaLysGluThrSerGluLysPheThrProAl	320	
QY				

QY	1240	-----	1241	-----	1242
Db	320	aLysGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMe	340		
QY	1241	-----	1242	-----	1243
Db	340	tSerProAlaIysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgGly	360		
QY	1261	GATCACATGGGAGGAAAAAACAATCTGTAAGACTGAATGCGTGCAGGAGTAACACC	1320		
Db	360	sIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrse	380		
QY	1321	TAATAAACTGAAGTTTTTGGAAAAAGAACATCTAATATGATTCATGTCCTCAAAAGA	1380		
Db	380	rAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGI	400		
QY	1381	AACATCTCAAAAAGCAAGTACAAATGTGGATGTGAGTCTCTGTAGAGCCTATATTCAGTCT	1440		
Db	400	uSerSerThrLysAlaSerAlaAsn	420		
QY	1441	TTTTGGGCACGGACTATTGAAAATTACAGTGTACAAAAGTTGAGGAACATTTAATCT	1500		
Db	408	-----	409	-----	410
QY	1501	TGCTACCAAGATTCTCTAAGAGTGCTGCACAGAAATTATACGTGTTTACCTGATGCTAC	1560		
Db	408	-----	409	-----	410
QY	1561	ATATCAAAAAGATATCAAAAACAATAATCAAAAATAGAAGATCAGATGTTCCCATCAGA	1620		
Db	409	-----	410	-----	411
QY	1621	ATCCAAACGAGAGCAAGATCAAGAAATATTCTTGGGATTCGGGAGTCTCTTCAGAGTTC	1680		
Db	415	uSerLysGlnGluLysAspGluLysSerCysAspSerArgSerLeuPheGluSerse	435		
QY	1681	TGCAAGACTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAATGGAGATAAATAG	1740		
Db	435	rAlaLysIleGlnValCysIleProGluSerIleTyrglnLysValMetGluIleAsnAr	455		
QY	1741	AGAAGTAGAGAGCTTCCTCAGAGGCATCTGCTCTCAAGCCTGCCGTGCAAAATGCAAAA	1800		
Db	455	gGluValGluGluProLysLysProSerAlaPheLysProAlaIleGluMetGlnAs	475		
QY	1801	GACTGTTCCTCAATAAAGCCTTTGAATTGAGAATGAACAAACATTCAGAGCAGCTCAGAT	1860		
Db	475	nSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMe	495		
QY	1861	GTTCCTCATCAGATCCAAACAAAGACGATGAAGAAAATCTTGGGATCTCAGAGTCC	1920		
Db	495	tPheProGluSerLysGlnLysAspTyrgluGluAsnSerTrpAspSerGluSerLe	515		
QY	1921	CTGTGACAGCTTTCAGAGAGATGTGATTATTACCCAAAGCTACACATCAAAAAGAAAT	1980		
Db	515	uCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIl	535		
QY	1981	CGATACCTTAAGTGGAAAATTAGAAGAGTCTCCTGTTAAAGATGGTCTCTGAAGCCTAC	2040		
Db	535	eAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaTh	555		
QY	2041	CTGTGGAAGAAAAGTTCTCTTCCAAATAAGCCTTAGAATTAAGACACAGAGAAACATT	2100		
Db	555	rCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPh	575		
QY	2101	CAAAAGCAGAGTCTCCTGATAAAGATGGTCTTCTGAAGCCTACTCTGTGGAAGGAAAGTTTC	2160		
Db	575	eLysAlaGluProGlyLysProSerAlaPheGluPro	598		
QY	2161	TC TTCCTCAATAAGCCCTAGAAATTAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGA	2220		
Db	589	-----AlaThrGluMetGlnLys	594		
QY	2221	TAATGATGGTCTTCTGAAGCCTACCTGTGGAAGGAAAGTTTCTTCTTCCAAATAAGCTTT	2280		



Db 595 -----SerValProAsnLysAlaLe 601  
 QY 2281 AGAATTGAAGGACAGAGAAACATTCAAACGAGCTCAGATGTTCCCATCAGAAATCAAAACA 2340  
 Db 601 uGluLeuLysAsnGluInThrTrpArgAlaAspGluLeuLeuProSerGluSerLysG 621  
 QY 2341 AAAGATGATGAAGAAATCTCTGGGATTTTCAGAGTTTCCTTGAGATCTCTTACAGAA 2400  
 Db 621 nLysAspTyrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLy 641  
 QY 2401 TGATGTGCTTTACCCCAAGCTTACACATCAAAAGAAATTCGATCTTAAGTGGAAATTT 2460  
 Db 641 sAspValCysLeuProLysAlaAlaHisGlnLysGluLeuAspLysIleAsnGlyLysLe 661  
 QY 2461 AGAAGAGTCTCTGTATAAGATGCTCTCTGAAGCTTACCTGTGAATCAAAATTTCTCT 2520  
 Db 661 uGluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerI 681  
 QY 2521 TCCAAATAAGCTTAGAATTGAAGGACAGAGAAACATTCAAGACAGAGGATGTCAGTTC 2580  
 Db 681 eProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLy 701  
 QY 2581 TGTAGAGTCCACATTCAGTCTTTTGGCAAAACCGACTACTGAAATTCACAGTCTACAAA 2640  
 Db 701 s----ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal-- 715  
 QY 2641 AGTGAGGAAGACTTAACTTACTTACCAAGGAGGAGCAACAAGACAGTAACCTGGACA 2700  
 Db 716 -----ProAsnLysAlaLeuGluLeuLy 723  
 QY 2701 ACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAGATCAACAATAAGATGCC 2760  
 Db 723 sAsnGluInThrLeuArgAlaAspGluLeuPro----- 735  
 QY 2761 CACATCAGAAATTAGAGAAAGAAAGATACAAATCAACTTCAGATTCTGAGATTAATCTC 2820  
 Db 736 ----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCy 754  
 QY 2821 TGTGAGTGATACACAGAAATTAGCTGTTTACCTGAGCTACATATCAAAAGAAATAAA 2880  
 Db 754 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluLeuAs 774  
 QY 2881 GACAAACAAATGCAAAATAGAAGAGTCT----- 2908  
 Db 774 physIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCy 794  
 QY 2908 ----- 2908  
 Db 794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLy 814  
 QY 2909 -----CCTGAAAGCCCTTCTCACTTTGAGCTGCGCACTGAAATGCAAACTCTGT 2958  
 Db 814 sAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834  
 QY 2959 TCCAAATAAGGCTTAGAATGAAGAATAAAACAAATTTGAGACA----- 3004  
 Db 834 lProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGlnMetPhePr 854  
 QY 3004 ----- 3004  
 Db 854 oSerGluSerLysGlnLysValGluGluAsnSerTrpAspSerGluSerLeuArgG 874  
 QY 3004 ----- 3004  
 Db 874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894  
 QY 3005 -----GATTCAACTACCTATCAAAATCTTGGATGCACTTCCTTC 3045  
 Db 894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914  
 QY 3046 TTGTGAAAGAGGAGGAACTTAAAAAGATAACTGTGAACAAATTCAGACAAATGGA 3105

Db 914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetG 934  
 QY 3106 ACAAATGAAAAATAAGTTTGTCTACTACAAAAAGGAACCTCTCAAGACGCAAGAAATAAA 3165  
 Db 934 uGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954  
 QY 3166 ATCAAGTTAGAGAACCAAAAAAGCTAAATCGGACAAAGAGCTCTGCAGTGTGAGATTGCC 3225  
 Db 954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuTh 974  
 QY 3226 TTTAAATCAAGAAAGAGAGAGAAATGTCGATATATTAAAGAAAAAATTAGACC 3285  
 Db 974 rLeuAsnGlnGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysIleArgG 994  
 QY 3286 C-----GAAGAGCAACTTAGCAAAAAAGTTAGAAAGTGAAGTGAACCAACTTGA 3330  
 Db 994 uGluLeuGlyArgIleGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuG 1014  
 QY 3331 ACAGACTCTCAGATACAAGATATAGAAATGAAAGTGTAAAGTAATTTGAATCAGGT 3390  
 Db 1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034  
 QY 3391 TTCTCACACTCATGAAGTGAATAATGATCTCTTTTCATGAAAAATTCATGTTGAAAAAGGA 3450  
 Db 1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysG 1054  
 QY 3451 AATTGCCATCTAAACTGGAAGTACCACTGCAACATCAACACAGCTGAGGAGAAA 3510  
 Db 1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGlnLysGluAs 1074  
 QY 3511 TAAATACCTTTGAGACATTAAGATTTTACAAGAAAAAGATGCTCAACTTCAAAATGACCCCT 3570  
 Db 1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094  
 QY 3571 AAACTGAAACAGAAACACAGTAAACAAAGCGGATCTCAGTATAGAGACGAGCTTAAAGT 3630  
 Db 1094 uLysLeuLysGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114  
 QY 3631 TCTGACGGCAGAGAACACGATGCTGACTTCTTAAATTCAGAGAA 3673  
 Db 1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128  
 RESULT 6  
 ABJ37784  
 ID ABJ37784 standard; protein; 1341 AA.  
 XX  
 AC ABJ37784;  
 XX  
 DT 15-MAY-2003 (first entry)  
 XX  
 DE Human tumour-related protein - SEQ ID No 565.  
 XX  
 DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;  
 KW tumour; breast cancer; cancer; immune response stimulation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283956-A1.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 15-APR-2002; 2002WO-US012378.  
 XX  
 PR 13-APR-2001; 2001US-00834759.  
 PR 07-DEC-2001; 2001US-00007805.  
 PR 13-FEB-2002; 2002US-00076622.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;  
 PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;  
 PI Vedvick TS, McNeill PD, Durham M;  
 XX

DR WPI; 2003-103376/09.  
 XX New polypeptide and polynucleotide useful for stimulating and/or  
 PT expanding T cells specific for a tumor protein and treating breast  
 PT cancer.  
 XX Example 9; Page 342-346; 375pp; English.  
 XX The invention comprises a method of stimulating and/or expanding T cells  
 CC specific for a tumor protein. The invention further comprises human  
 CC nucleic acids and proteins that are associated with tumors (e.g. breast  
 CC cancer). The method and sequences of the invention are useful for:  
 CC stimulating and/or expanding T cells specific for a tumor protein,  
 CC detecting the presence of cancer, stimulating an immune response in a  
 CC patient and treating breast cancer. The present amino acid sequence  
 CC represents a human tumour-related protein  
 XX Sequence 1341 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,38e-273 Length: 1341  
 Score: 3458.00 Matches: 740  
 Percent Similarity: 65.58% Conservative: 83  
 Best Local Similarity: 58.96% Mismatches: 159  
 Query Match: 53.63% Indels: 273  
 DB: 6 Gaps: 12  
 US-09-602-362E-26 (1-3673) x ABJ37784 (1-1341)  
 QY 349 ATGACAGTGGGAAGAGCCCTCAACCTGACGACAAAGAGATGATGAAGAGGAGCTGCT 408  
 DB 1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20  
 QY 409 CTACACTGGCCCTGTCTCAATGCCATGCCATGCAAGTAGTACATTTCTGGTACAGAGAAG 468  
 DB 21 LeuHisThrAlaCysValAsnGlyHisGluValValThrPheLeuValAspArgLys 40  
 QY 469 TGCNCGCTTAATCTCTGATGCGAAGGAGGACACCTCTCATGAAGCTCTCAACAATGC 528  
 DB 41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60  
 QY 529 GAGAGGAAGCTTTGTGCAAAATATCTCATAGCTGGTGTCTGATCTCTAAATATGTAGA 588  
 DB 61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80  
 QY 589 TGTGTATGCAACACCGCTCTCATATGCTGCTTATAGTGAGAAATTTATTAATGCTGCG 648  
 DB 80 pValThrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAl 100  
 QY 649 AACACTGCTGCTCTATGCTGAGTATCGAGGTGCAAAACAGGCTAGCCTCACACCCCT 708  
 DB 100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120  
 QY 709 TTTACTGGCCATACAGAAAGAGCAAGCAACTGTGGAATTTTACTAACCAAAAATGC 768  
 DB 120 uLeuLeuSerIleThrLysArgSerGlnIleValGluPheLeuLeuIleLysAsnAl 140  
 QY 769 AAATGCAACCGCATTTAATGAGTCTAAATGCAAGCCCTCATGCTGGCATATGTGAAGG 828  
 DB 140 AsnAlaAsnAlaValAsnLysTyrLysCysThrAlaLeuMetLeuAlaValCysHisGln 160  
 QY 829 CTCATCAGAGATAGTGGCATCTCTTTCAGCAAAATGTTGACGTCTTGTGTAAGACAT 888  
 DB 160 ySerSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIl 180  
 QY 889 ACATGGAATAACTGCAGACAGTTATCTGCTGCTGCTGAGTAAATTCATTCATCAACA 948  
 DB 180 eCysGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGln 200  
 QY 949 ACTTTTGGACATATACGAAAATTTACTAAAATCTCTCAAAATACCAATCCAGAGGAAAC 1008  
 DB 200 nileMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyTh 220

QY 1009 ATCTACAGGAACACCTGATGAGGTGCACCCCTTGGCG- 1045  
 DB rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGln 240  
 QY 1046 -----GAAAGAACACCTGCAC 1062  
 DB 240 uSerLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspTh 260  
 QY 1063 GCGTGAAGCTTGTGCAAAAAACACCTGACAGGCTGCACGCTTGGTGAGGAAACGTC 1122  
 DB 260 rAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSe 280  
 QY 1123 TGCACAAATTCATGTCTGGGAAAGCAACATCTGAAAGTTTGAACAGTCAACACAGA 1182  
 DB 280 rAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGln 300  
 QY 1183 AACACCTAGGAAAAATTTGAGGCTTACAAAGAAACATCTGAGAAATTTTCATGGCCA- 1240  
 DB 300 uThrProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAl 320  
 QY 1240 ----- 1240  
 DB 320 aLysGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMe 340  
 QY 1241 -----GCAAAAGAAAGATCTAGGAA 1260  
 DB 340 tSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLys 360  
 QY 1261 GATCACATGGAGGAAAAAGAAACATCTGTAAGACTGAATGCGTGGCGAGGATACACC 1320  
 DB 360 sIleAlaTrpGluLysGluThrProValLysThrGlyCysValAlaArgValThrSe 380  
 QY 1321 TATAAACTCGAGTTTGGAAAAAGCAACATCTAATATGATGTCATGTCCTCAACAAAGA 1380  
 DB 380 rAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGln 400  
 QY 1381 AACATCTACAAAAAGCAAGTCAAAATGTGGATGTGAGTCTCTGTAGAGCCTATATTCAGTCT 1440  
 DB 400 uSerSerThrLysAlaSerAlaAsn----- 408  
 QY 1441 TTTTGGCACACGGACTATTGCAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCT 1500  
 DB 408 ----- 408  
 QY 1501 TGCTACCAAGATATTCTTAAGAGTGTGCACAGAATTATACGTGTTTACCTGATGCTAC 1560  
 DB 408 ----- 408  
 QY 1561 ATATCAAAAAGATATCAAAAATAATCAAAAATAGAAAGATCAGATGTTCCCATCAGA 1620  
 DB 409 -----AspGlnArgPheProSerGln 415  
 QY 1621 ATCCAAACGAGAGCAGATCAAGAAATATTCTTGGGATCTCTGGAGTCTCTTTTGGAGTTC 1680  
 DB 415 uSerLysGlnGluLysGluLysArgSerCysAspSerArgSerLeuPheGluSerSe 435  
 QY 1681 TGCACAGCTCAAGTGTGTATACCTGAGTCTATGTATCAGAAAGTAAATGAGATAATAG 1740  
 DB 435 rAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnAr 455  
 QY 1741 AGAAGTAGAAGAGCTTCTGAGAGCCATCTGCTTCAAGCCCTGCGCTGAAATGCAAAA 1800  
 DB 455 gGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAs 475  
 QY 1801 GACTGTTTCCAAATAAGCCCTTTCAATGAAGATGAACAAACATTGAGAGCAGCTCAGAT 1860  
 DB 475 nSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMe 495  
 QY 1861 GTTCCCATCAGAAATCAAAACAAAGGACGATGAAGAAAATCTTGGGATTTCTGAGAGTCC 1920  
 DB 495 tPheProProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLe 515  
 QY 1921 CTGTGACAGCGTTTTCAGAGAGGATGTGATTATTTACCCCAAGCTTACACATCAAAAGAAAT 1980

```

515 uCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluI 535
1981 CGATACCTTAAGTGGAAATTAAGAGAGCTCTCTGTTAAAGATGCTTCTGAAGCCTAC 2040
535 eAspLysIleAsnGlyLysLeuGluSerProAsnLysAspGlyLeuLeuLysAlaTh 555
2041 CTGTGGAAGCAAGTTTCTTCCAAATAAAGCCTTAGAATTAAGGACAGAGAACACATT 2100
555 rCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPh 575
2101 CAAAGCAGAGTCTCTGTAATAAGATGGTCTTCTGAAGCCTACCTGTGGAGGAAAGTTTC 2160
575 eLysAlaGluProGlyLysProSerAlaPheGluPro----- 588
2161 TCTTCCAAATAAGCCTTAGAATTAAGGACAGAGAAACACTCAAAGCAGAGTCTCCTGA 2220
589 -----AlaThrGluMetGlnLys----- 594
2221 TAATGATGCTCTTCTGAAGCCTACCTGTGGAGGAAAGTTTCTTCCAAATAAGCTTT 2280
595 -----SerValProAsnLysAlaLe 601
2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAAATCCAAACA 2340
601 uGluLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysG 621
2341 AAAGATGATGAAGAAATTTCTGGATTCTGAGAGTTTCCCTGAGACTCTCTTACAGAA 2400
621 nLysAspTyrGluGluAsnSerTrpAspThrGluSerLysGluThrValSerGlnLys 641
2401 TGATGTGTTTACCCAGGCTACACATCAAAGAAATTCGATACCTTAAGTGGAAATTT 2460
641 sAspValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLe 661
2461 AGAAGAGTCTCTGTATAAGATGGTCTCTGAAGCCTACCTGTGAATGAATAATTTCTCT 2520
661 uGlyGlySerProValLysAspGlyLeuLeuLysAlaAsnGlyLysValSerIle 681
2521 TCCAAATAAGCCTTAGAATTAAGGACAGAGAAACATTCAAAGCAGAGGATGTGAGTTC 2580
681 eProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProGluLys 701
2581 TGTAGATCCACATTACAGTCTTTTGGCAACCGACTACTGAAATTCACAGTCTACAAA 2640
701 s----ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal-- 715
2641 AGTTGAGGAAGACTTTAATCTTACTACCAAGGAGGAGCAACAAAGCAGACTGAGGACA 2700
716 -----ProAsnLysAlaLeuGluLeuLys 723
2701 ACAGGACGTGATATTGGCATATTGAACGAGCTCCACAGATCAACAAATAAGATGCC 2760
723 sAsnGluGlnThrLeuArgAlaAspGluIleLeuPro----- 735
2761 CACATCAGAAATTAAGGAGAAAGAGATACAAATCAACTTCAGATTTCTGAGATTATCTC 2820
736 ----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCy 754
2821 TGTGATGATACACAGAAATTATGAGTGTACCTGAGGCTCATATCAAAAAGAAATAAA 2880
754 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 774
2881 GACACAAATGCAAAATAGAGACTCT----- 2908
774 pLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCy 794
2908 ----- 2908
794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 814
2909 -----CCTGAAAGCCTTCTCACTTTGAGCGCTGCCACTGAAATGCAAAACTCTGT 2958

```

```

814 sAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834
2959 TCCAAATAAGGCTTAGAATGGAGAAATAACAACATTTAGAGACA----- 3004
834 lProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854
3004 ----- 3004
854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgG 874
3004 ----- 3004
874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894
3005 -----GATTCAACTACCTTACCAATCAAAATCTTGGATGCACCTTCCTTC 3045
894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914
3046 TTGTGAAGAGGAGGAACTTAAAGAAATAACTGTGAACAAATTAACAGCAAAATTCGA 3105
914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetG 934
3106 ACAATCAAAAATAAGTTTGTCTACTACAAAAGAACTGTCAAGACCGAAGAAATAAA 3165
934 uGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLy 954
3166 ATCAGACTTAGAGAACCAAAAGCTAAATGGCAACAGAGCTCTGCACTGTGAGATTGCC 3225
954 sSerGlnLeuGluAsnGlnLysValLysTyrGluGlnGluLeuCysSerValargLeuTh 974
3226 TTAAATCAAGAGAGAGAGAGAAATGTCGATATATTAAGAAAAATTAAGAC 3285
974 rLeuAsnGlnGluGluLysArgArgAlaAspIleLeuAsnGluLysIleArgG 994
3286 C-----CAAGAGCAACTTAGCAAAAAGTTAGAGTGAACACCAACTTGA 3330
994 uGluLeuGlyArgIleGluGlnGlnHisArgLysGluLeuGluValLysGlnGlnLeuG 1014
3331 ACAGACTCTCAGATACAGATATAGAAATTAAGTGAACAGTGAACAGTAAATTTGAATCAGGT 3390
1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVa 1034
3391 TTCTCACACTCAGTAAAGTCAAAATGATCTCTTCATGAAAATTCGATGTTGAAAAAGGA 3450
1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysG 1054
3451 AATTGCCATCTAAAACTGGAAGTAGCCACACTGAAACATCAACACAGGTGAAGGAAAA 3510
1054 uIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074
3511 TAAATACCTTTGAGGACATTAAGATTTTACAGAAAGAAATGCTGAACCTTCAATGACCCT 3570
1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094
3571 AAAACTGAAACAGAAACAGATAACAAAAGGCGATCTCAGTATAGAGAGCGCTTAAAGT 3630
1094 uLysLeuLysGluLeuSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114
3631 TCTGACGCGCAGAGAACACGATGCTGACTTCTAAATTTGAAGGAA 3673
1114 lLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128
RESULT 7
ABR47548
ID ABR47548 standard; protein; 1341 AA.
XX
AC ABR47548;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated protein sequence SEQ ID NO:334.
XX
KW Human; breast cancer; cytostatic; gene therapy.

```

XX OS Homo sapiens.  
 XX PN WO2003004989-A2.  
 XX PD 16-JAN-2003.  
 XX PF 21-JUN-2002; 2002WO-US019669.  
 XX PR 21-JUN-2001; 2001US-0299887P.  
 XX PR 27-JUN-2001; 2001US-0301572P.  
 XX PR 18-JUL-2001; 2001US-0306501P.  
 XX PR 25-SEP-2001; 2001US-0325002P.  
 XX PR 05-MAR-2002; 2002US-0362585P.  
 XX PR 14-MAY-2002; 2002US-0380391P.  
 XX PA (MILL-) MILLENIUM PHARM INC.  
 XX PI Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;  
 XX PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;  
 XX PI Bast RC, Hortobagyi GN, Pusztai L, Meric P, Sahin A, Mills GB;  
 XX PI WPI; 2003-210381/20.  
 XX DR N-PSDB; ACC50246.  
 XX XX  
 XX PT Breast cancer diagnosis or treatment by comparing the level of expression  
 XX PT of a marker in a patient sample with that in the control non-breast  
 XX PT cancer sample.  
 XX PS Claim 1; SEQ ID NO 334; 128pp; English.  
 XX CC The present invention describes a method for assessing whether a patient  
 XX CC is afflicted with breast cancer. The method comprises comparing the level  
 XX CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and  
 XX CC ABR47386 to ABR47632) in a patient sample and the normal level of  
 XX CC expression of the marker in a control non-breast cancer sample, where a  
 XX CC significant increase in the level of expression of the marker in the  
 XX CC patient sample and the normal level is an indication that the patient is  
 XX CC afflicted with breast cancer. The breast cancer associated sequences from  
 XX CC the present invention have cytostatic activities and can be used in gene  
 XX CC therapy. The method is useful for diagnosing and treating breast cancer.  
 XX CC N.B. The sequence data for this patent did not form part of the printed  
 XX CC specification, but was obtained in electronic format directly from WIPO  
 XX CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 1341 AA;

Alignment Scores:  
 Pred. No.: 2,38e-273 Length: 1341  
 Score: 3458.00 Matches: 740  
 Percent Similarity: 65.58% Conservative: 83  
 Best Local Similarity: 58.96% Mismatches: 159  
 Query Match: 53.63% Indels: 273  
 DB: 6 Gaps: 12

US-09-602-362E-26 (1-3673) x ABR47548 (1-1341)

QY 349 ATGACAGTAGGGAAGAGCCGCTCACTGACAAAGAGATATGAGGAAGAGGACTGCT 408  
 Db 1 MetThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAla 20  
 QY 409 CTACACTGGCCGTGTGTCATGCGCCATGCANAGTAGTAACTATTCCTGCTAGACAGAAAG 468  
 Db 21 LeuHisTrpAlaCysValaAsnGlyHisGluGluValValThrPheLeuValAspArgLys 40  
 QY 469 TGCNGCTTAATCTCTTGTATGCGGAGGAGGACACCTCTGATGAAGGCTCTTACATGCC 528  
 Db 41 CysGlnLeuAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCys 60  
 QY 529 GAGAGGAGACCTTTGTGCAATATCTCATAGATGCTGGTGTGATCTAAATATGTAGA 588  
 Db 61 HisGlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAs 80

QY 589 TGTGTATGGCAACACAGCGCTCTCCATTATGCGGTTTATAGTGAGAAATTTATTATGTGGC 648  
 Db 80 pValTyGlyAsnMetAlaLeuHisTyRAlaValTySerGluIleLeuSerValValAl 100  
 QY 649 AACACTGCTCTCTATGTCAGTCTCATCGAGGTGCAAAACAGGCTAGCTCACACCCCT 708  
 Db 100 aLysLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLe 120  
 QY 709 TTTACTGGCCATACAGAAAGCAAGCAACTGTGGAATTTTACTACAAAAATGC 768  
 Db 120 uLeuLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAl 140  
 QY 769 AAATGCAACGCAATTTAATGAGTCTAAATGCACAGCCCTCATCTGCGCATATGTGAAG 828  
 Db 140 aAsnAlaAsnAlaValAsnLysTyRLeuValHisGlnMetLeuAlaValCysHisG 160  
 QY 829 CTATCAGAGATAGTCGGCATGCTTCTTCAGCAAAATGTTGACGCTCTTGTGCAACACAT 888  
 Db 160 ySerSerGluIleValGlyMetLeuLeuGlnAsnValAspValPheAlaAlaAspI 180  
 QY 889 ACATGGAATAACTGCAGAACGTTTATGCTGCTGCTGCTGAGTTAATTAATCATCAACA 948  
 Db 180 eCysGlyValThrAlaGluHisTyRAlaValThrCysGlyPheHisIleHisGlu 200  
 QY 949 ACTTTTGGACATATACGAAATTTACTTAAATCTCAAAATACCAATCCAGATCCAGAGAAC 1008  
 Db 200 nIleMetGluTyIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGly 220  
 QY 1009 ATCTACAGGAACACCTGATGAGGCTGCACCCCTGGGCG- 1045  
 Db 220 rSerAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAla 240  
 QY 1046 -----GAAAGAACCTCACAC 1062  
 Db 240 uSerLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAsp 260  
 QY 1063 GGCTGAAAGCTTCTGGAAAAACACCTGACGAGGCTGCACGCTTGTGTGAGGGAACGTC 1122  
 Db 260 rAlaGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThr 280  
 QY 1123 TGCACAAATTCATGCTCGGGAAAGCAACATCTGGAAGCTTGAACAGTCAACAGAGA 1182  
 Db 280 rAspLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGlu 300  
 QY 1183 AACACCTAGGAAATTTTGAGGCTACAAAGAACATCTGAGAAATTTTCATGGCCA- 1240  
 Db 300 uThrProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAl 320  
 QY 1240 ----- 1240  
 Db 320 aLysGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMe 340  
 QY 1241 -----GCAAAAGAAAGTCTAGGA 1260  
 Db 340 tSerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArg 360  
 QY 1261 GATCACATGGGAGGAAAAAGAAACATCTGTAAGACACTGAATGCTGGCAGGAGTAACACC 1320  
 Db 360 sIleAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSe 380  
 QY 1321 TAATAAACCTGAAGTTTTGGAAAAAGGAACATCTAATATGATTCATGCTGCTCACAAGA 1380  
 Db 380 rAsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysG 400  
 QY 1381 AACATCTACAAAACAGTACAAATGCGATGTGAGTCTCTGTAGAGCCTATATTCTAGTCT 1440  
 Db 400 uSerSerThrLysAlaSerAlaAsn----- 408  
 QY 1441 TTTTGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTTATCT 1500  
 Db 408 ----- 408  
 QY 1501 TGCTACCAAGATTATCTTAAGAGTGTGCACAGAAATTATACGGTTTACCTGTGTCTAC 1560

Db 408 ----- 408  
QY 1561 ATATCAAAAGATATCAAAACCAATAAATCACAAATAGAAAGATCAGATGTTCCCATCAGA 1620  
Db 409 ----- 409  
QY 1621 ATCCAAACGAGGAGAGATGAAGAATATTCTTGGGATTCGGAGTCTCTTTGAGAGTTTC 1680  
Db 415 uSerLysGlnGluGluuAspGluGluTyrSerCysAspSerAArgSerLeuPheGluSerSe 435  
QY 1681 TGCAAAAGACTCAAGTGTGTATACCTCAGTCTATGATCAGAAAGTAATGAGATATAAATAG 1740  
Db 435 rAlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluLileAsnAr 455  
QY 1741 AGAAGTAGAAGACTCTTCGAGAGCCATCTGCCTTCAAGCTGCGGTNGAAATCAAAA 1800  
Db 455 gGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAs 475  
QY 1801 GACTGTTCCTCAAAATAAGCTTTGAATTGAAGTAATGAACAAACATTTGAGAGCAGCTCAGAT 1860  
Db 475 nSerValProAsnLysAlaPheGluLysAsnGluGlnThrLeuArgAlaAspProWe 495  
QY 1861 GTTCCCATCAGATCCAAACAAAGAGCAGATGAAGAAATTTCTGGGATTTCTGAGAGTCC 1920  
Db 495 tPheProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSerLe 515  
QY 1921 CNGTGAGCGGTTTCACAGAGGATCTGTATTACCCAAAGCTACACATCAAAAGATTT 1980  
Db 515 uCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIl 535  
QY 1981 CGATACCTTTAAGTGGAAATTAAGAAGTCTCCTGTTAAAGATGCTCTTCAAGCTTAC 2040  
Db 535 eAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaTh 555  
QY 2041 CTGTCGAGGAAAGTTTCTCTCCAAATTAAGCCTTAGAATTAAGGACAGAGAAACATT 2100  
Db 555 rCysGlyMetLysValSerIleProThrLysAlaLeuGluLysAspMetGlnThrPh 575  
QY 2101 CAAAGCAGAGTCTCCTGTATAAGTGGTCTTCTGAAGCTACCTGTGGAAGGAAGTTTC 2160  
Db 575 eLysAlaGluProProGluLysProSerAlaPheGluPro 588  
QY 2161 TCTTCCAAATAAGCCTTAGAATTAAGGACAGAGAAACACTCAAGACAGAGTCTCCTGA 2220  
Db 589 -----AlaThrGluMetGlnLys----- 594  
QY 2221 TAATGATGCTCTCTGAAGCCTTACCTGTGGAGGAAGTTTCTCTTCCAAATAAGCTTT 2280  
Db 595 -----SerValProAsnLysAlaLe 601  
QY 2281 AGAATTGAAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAAATCCAAACA 2340  
Db 601 uGluLysAsnGluGlnThrTrpArgAlaAspGluLileLeuProSerGluSerLysGl 621  
QY 2341 AAAGATGATGAAGAAATTTCTGGATTTTGAGAGTTTCTCTGAGACTCTCTTACAGAA 2400  
Db 621 nLysAspTyrGluGluAsnSerTrpAspThrGluSerLysCysGluThrValSerGlnLy 641  
QY 2401 TGATGTGTGTTTACCACAGGCTACACATCAAAAAGAAATTCGATACCTTAAGTGAAATTT 2460  
Db 641 sAspValCysLeuProLysAlaAlaHisGlnLysGluLileAspLysIleAsnGlyLysLe 661  
QY 2461 AGAAGAGTCTCTGTATAAGATGTTCTTCTGAAGCCTTACCTGTGGAATGAATAATTTCTCT 2520  
Db 661 uGluGlySerProValLysAspGlyLeuLysAlaAsnCysGlyMetLysValSerIl 681  
QY 2521 TCCAAATAAGCCTTAGAATTAAGGACAGAGAAACATTCAAGCAGAGGATGAGTTTC 2580  
Db 681 eProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProGluLy 701  
QY 2581 TGTAGAGTCCACATTACAGTCTCTTTTGGCAAAACCGACTACTGAAATTCACAGTCTACAAA 2640  
Db ----- 2640

Db 701 s---ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal-- 715  
QY 2641 AGTTGAGGAAGACTTTAATCTTACTTACCAAGGAGGAGCAACAAAGACAGTAACTGGACA 2700  
Db 716 -----ProAsnLysAlaLeuGluLeuLy 723  
QY 2701 ACAGAAAGCTGATATTGGCATTTATTGAACGAGCTCCCAAGATCAAAACAAATAAGATGCC 2760  
Db 723 sAsnGluGlnThrLeuArgAlaAspGluLileLeuPro----- 735  
QY 2761 CACATCAGAAATTAGGAAGAAAGAGATACAAATCAACTTCAGATTCTGAGATTATCTC 2820  
Db 736 -----SertGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCy 754  
QY 2821 TGTGATGATACACAGAAATTTATGAGTCTTTACCTGAGGCTACATATCAAAAGAAATAAAA 2880  
Db 754 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluLileAs 774  
QY 2881 GACAAACAATGGCAAAATAGAGAGTCT----- 2908  
Db 774 pLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCy 794  
QY 2908 ----- 2908  
Db 794 sArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLy 814  
QY 2909 -----CCTGAAAGGCTTCTCACTTTGAGCCTGCCACTGAAATGCAAAACTCTGT 2958  
Db 814 sAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVa 834  
QY 2959 TCCAAATAAGGCTTAGAATGGAAGATAAACAACAACTTCAGAGCA----- 3004  
Db 834 lProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePr 854  
QY 3004 ----- 3004  
Db 854 oSerGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGl 874  
QY 3004 ----- 3004  
Db 874 uThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLy 894  
QY 3005 -----GATTCAACTACCTTATCAAAATCTTGATGCACTCTCTTC 3045  
Db 894 sIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSe 914  
QY 3046 TTCTGAAAGGAGGAGGAACCTTAAAGATAACTGTGTAACAATTTACAGCAAAATGGA 3105  
Db 914 rCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGl 934  
QY 3106 ACAAATCAAAAAATAAGTTTGTGTACTACAAAAGGAACCTGTCAAGACCGAAAGAAATAAA 3165  
Db 934 uGlnMetLysLysLysPheCysValLeuLysLysLysLysLeuSerGluAlaLysGluLileLy 954  
QY 3166 ATCAGAGTTAGAGAACCAAAAGCTAAATGGGAAACAGAGCTCTGAGTGTGAGATTGCC 3225  
Db 954 sSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuLysCysSerValArgLeuth 974  
QY 3226 TTTAAATCAAGAGAGAGAGAGAGAAATGTCGATATATTAAAAGAAAAAATTAGACC 3285  
Db 974 rLeuAsnGlnGluGluGluLysArgArgAsnAlaAspIleLeuAsnGluLysLileArgGl 994  
QY 3286 C-----GAAGACCACTTAGGAAAAAGTTAGAAAGTGAACACCAACTTGA 3330  
Db 994 uGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGl 1014  
QY 3331 ACAGACTCTCAGATACAGATATAGATTAAGAGTGAACAGTAACTAATTTGAATCAGGT 3390  
Db 1014 uGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLysGlnVa 1034  
QY 3391 TTCTCACTCATGAAAGTGAATGATCTCTTTTATGAAAAATTCATCTTCTGAAAAAGGA 3450  
Db 1034 lSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGl 1054

QY 3451 RATTGCGATGCTAAACTGGAAGTAGCCACACTGAAACATCAACACAGGTGAAGGAAA 3510  
 Db 1054 uilealaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAs 1074  
 QY 3511 TAAATPACTTTGAGGACATTAAGATTTTACAGAAAAGAAATGCTCAACTTCAAAATGACCCCT 3570  
 Db 1074 nLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLe 1094  
 QY 3571 AAAACCTGAAACAGAAACAGTAACAAAGAGGCACTCTCATATAGAGAGAGAGCTTAAAGT 3630  
 Db 1094 uLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVa 1114  
 QY 3631 TGTGAGGCGAGAGAACAGCTGCTCACTTCTTAAATTGAAGGAA 3673  
 Db 1114 lleuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 1128

RESULT 8

ID ABJ37788 standard; protein; 1349 AA.

XX ABJ37788;

DT 15-MAY-2003 (first entry)

XX Human tumour-related protein - SEQ ID No 573.

DE Human; vaccines; gene therapy; T cell stimulation; T cell expansion;  
 KW tumour; breast cancer; cancer; immune response stimulation.

XX Homo sapiens.

PN WO200283956-A1.

XX 24-OCT-2002.

PF 15-APR-2002; 2002WO-US012378.

XX 13-APR-2001; 2001US-00834759.

PR 07-DEC-2001; 2001US-00007805.

PR 13-FEB-2002; 2002US-00076622.

XX (CORI-) CORIXA CORP.

XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;  
 PI Mitcham JL, Xu J, Hazlock SL, Hepler WT, Henderson RA, Fanger CR;  
 PI Vedvick TS, McNeill PD, Durham M;

XX WPI; 2003-103376/09.

XX New polypeptide and polynucleotide useful for stimulating and/or  
 PT expanding T cells specific for a tumor protein and treating breast  
 PT cancer.

XX Example 12; Page 353-357; 375pp; English.

XX The invention comprises a method of stimulating and/or expanding T cells  
 CC specific for a tumour protein. The invention further comprises human  
 CC nucleic acids and proteins that are associated with tumours (e.g. breast  
 CC cancer). The method and sequences of the invention are useful for  
 CC stimulating and/or expanding T cells specific for a tumour protein,  
 CC detecting the presence of cancer, stimulating an immune response in a  
 CC patient and treating breast cancer. The present amino acid sequence  
 CC represents a human tumour-related protein

XX Sequence 1349 AA;

Alignment Scores:

Pred. No.: 7.07e-272 Length: 1349  
 Score: 3440.00 Matches: 737  
 Percent Similarity: 65.39% Conservative: 83  
 Best Local Similarity: 58.77% Mismatches: 161  
 Query Match: 53.35% Indels: 273

DB: 6 Gaps: 12  
 US-09-602-362E-26 (1-3673) x ABJ37788 (1-1349)  
 QY 352 ACAGTGGGAAGAGCCCGTCAACCTGAAACAAAGAGATATGAAGAAGAGGAGCTCTCTA 411  
 Db 10 ThrLysArgLysLysThrIleAsnLeuAsnIleGlnAspAlaGlnLysArgThrAlaLeu 29  
 QY 412 CACTGGGCTCTGCTCAATGCGCCATGCANAAAGTAGTACATTTCTGTTAGACAGAAAGTC 471  
 Db 30 HistpAlaCysValAsnGlyHisGluGluValValThrPheLeuValAspArgLysCys 49  
 QY 472 CNGCTTAATCTCCCTGATGCGGAAGGAGGACACCTCTCATGAAGGCTCTACAATGCCAG 531  
 Db 50 GlnProAspValLeuAspGlyGluHisArgThrProLeuMetLysAlaLeuGlnCysHis 69  
 QY 532 AGGGAAGCTTTGCAAAATATTCTCATAGATGCTGCTGCTCTAAATATTATGATGATGT 591  
 Db 70 GlnGluAla-CysAlaAsnIleLeuIleAspSerGlyAlaAspIleAsnLeuValAspVa 89  
 QY 592 GTATGGCAACACGGCTCTCCATTATGCCCTTTTATAGTCAGAAATTTTAAATGCTGGCAAC 651  
 Db 89 lTyrGlyAsnMetAlaLeuHisTyrAlaValTyrSerGluIleLeuSerValValAlaLys 109  
 QY 652 ACTGCTGCTCTATGCTGAGTCTCATGAGTGCAGGTCGCAAAACAGGCTGCTCACCCCTTT 711  
 Db 109 sLeuLeuSerHisGlyAlaValIleGluValHisAsnLysAlaSerLeuThrProLeuLe 129  
 QY 712 ACTGGCCATACAGAAAAGCAAGCAAACTCTGGAAATTTTACTTAACAAAAAATGCAAA 771  
 Db 129 uLeuSerIleThrLysArgSerGluGlnIleValGluPheLeuLeuIleLysAsnAlaAs 149  
 QY 772 TGCAAAAGCATTTAATGAGTCTAAATGACACAGCCCTCATGCTGCCATATGTGAAGGCTC 831  
 Db 149 nAlaAsnAlaValAsnLysLysCysThrAlaLeuMetLeuAlaValCysHisGlyLe 169  
 QY 832 ATCAGAGATAGTCGGCATGCTTCTTACGCAAAATCTTACGCTCTTTGCTGAAGACATACA 891  
 Db 169 uSerGluIleValGlyMetLeuLeuGlnGlnAsnValAspValPheAlaAlaAspIleCy 189  
 QY 892 TGAATAACTGCGAGACGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951  
 Db 189 sGlyValThrAlaGluHisTyrAlaValThrCysGlyPheHisHisIleHisGluGlnI 209  
 QY 952 TTGGAACATATACGAAAATTTACCTAAATAATCCTCAAAATACCAATCCAGAAAGAACATC 1011  
 Db 209 eMetGluTyrIleArgLysLeuSerLysAsnHisGlnAsnThrAsnProGluGlyThrSe 229  
 QY 1012 TACAGGAACACCTGATGAGGCTGCGCCCTTGGCG----- 1045  
 Db 229 rAlaGlyThrProAspGluAlaAlaProLeuAlaGluArgThrProAspThrAlaGluSe 249  
 QY 1046 -----GAAAGAACACCTGCACAGGC 1065  
 Db 249 rLeuValGluLysThrProAspGluAlaAlaProLeuValGluArgThrProAspThrAl 269  
 QY 1066 TGAAGCTTGTGGAAGAAACACCTGACAGGCTGCAGCTTGTGTTGGAGGAAAGTGTGC 1125  
 Db 269 aGluSerLeuValGluLysThrProAspGluAlaAlaSerLeuValGluGlyThrSerAs 289  
 QY 1126 CAAATTTCAATCTCTGGGAAAGCAACATCTGGAAGTTTGAACAGTCAACAGAAAGAAC 1185  
 Db 289 pLysIleGlnCysLeuGluLysAlaThrSerGlyLysPheGluGlnSerAlaGluGluTh 309  
 QY 1186 ACCTAGAAAATTTGAGCCCTTACAAAAGAACATCTGAGAAATTTTCATGCCCA----- 1240  
 Db 309 rProArgGluIleThrSerProAlaLysGluThrSerGluLysPheThrTrpProAlaL 329  
 QY 1240 ----- 1240  
 Db 329 sGlyArgProArgLysIleAlaTrpGluLysLysGluAspThrProArgGluIleMetSe 349  
 QY 1241 -----GCAAAAGAAAGATCTAGGAAGAT 1263

Db 349 rProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLysI 369  
QY 1264 CACATGGGAGGAAAAAGAACATCTGTAAAGACTGAATCGTGGCAGGAGTAACCTAA 1323  
Db 369 eAlaTrpGluLysLysGluThrProValLysThrGlyCysValAlaArgValThrSerAs 389  
QY 1324 TAAACTGAAGTTTTCGAAAAAGGAACATCTAATATGATGATGCTCTACAAAAAACA 1383  
Db 389 nLysThrLysValLysGluLysGlyArgSerLysMetIleAlaCysProThrLysGluSe 409  
QY 1384 ATCTCAAAAAGCAAGTACAAATGTGATGTGAGTCTGTAGAGCCTATATTCAGTCTTT 1443  
Db 409 rSerThrLysAlaSerAlaAsn 416  
QY 1444 TGGCACACGGACTATTGAAAATTCACAGTGTACAAAAGTTGAGGAAGACTTTAATCTTGC 1503  
Db 416 416  
QY 1504 TACCAAGATTATCTTAAGAGTGTGCACAGAAATTATACGTGTTTACCTGATGCTACATA 1563  
Db 416 416  
QY 1564 TCAAAAAGATATCAAAAATAAATACAAAATAGAAGATCAGATGTTCCCATCAGAAATC 1623  
Db 417 417  
QY 1624 CAAACGAGAGGAAGATGAAGATATTTCTGGGATTCCTGGGATCTCTTTGAGAGTTCTGC 1683  
Db 424 rLysGluGluLysGluLysGluLysSerCysAspSerArgSerLeuPheGluSerAl 444  
QY 1684 AAAGACTCAAGTGTGTATACCTGAGTCTATGATGATCAGAAAGTAATGGAGATAATAGAGA 1743  
Db 444 aLysIleGluValCysIleProGluSerIleThrGlnLysValMetGluIleAsnArgG 464  
QY 1744 ACTAGAAGACTCTCTGAGAACCCATCTGCTTCAAGCTCGCGTNGAAATGCAAAAGAC 1803  
Db 464 vAlGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsnSe 484  
QY 1804 TGTTCCAAATGAAGCTTTGAATTCAGAAATGAACATGCAACATTCAGAGAGCTCAGATGTT 1863  
Db 484 rValProAsnLysAlaPheGluLeuLysGluGluGlnThrLeuArgAlaAspProMetPh 504  
QY 1864 CCATCAGAAATCAAAACAAAGGAGCATGAAGAAATCTTGGGATTCGTAGAGTCCCTG 1923  
Db 504 eProGluSerLysGlnLysAspTyrGluLysAsnSerTrpAspSerGluSerLeuCy 524  
QY 1924 TGAGACGGTTTCACAGAGGATGTGTATTTACCCAAAGCTACATCAAAAAGAAATCGA 1983  
Db 524 sGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAs 544  
QY 1984 TACCTTAAGTGGAAATTAGAAGAGTCTCTGTTAAAGATGCTCTCTGAGCCTACCTG 2043  
Db 544 pLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAlaThrCy 564  
QY 2044 TGGAAAGGAAGTTCTCTTCCAAATAAAGCCTTAGAATTAAGGACAGAGAAACATTCAA 2103  
Db 564 sGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPheLy 584  
QY 2104 AGCAGAGTCTCTGATTAAGATGCTCTGAGCCTACCTGTGGAAGGAAGTTCTCT 2163  
Db 584 sAlaGluProProGlyLysProSerAlaPheGluPro 596  
QY 2164 TCCAAATAAGCCTTAGAATTAAGGACAGAGAAACACTCAAGAGAGTCTCCTGTATA 2223  
Db 597 597  
QY 2224 TGATGCTCTCTGAAGCCTACTCTGGAAGAAAGTTTCTCTTCCAAATAAAGCTTAGA 2283  
Db 603 603  
QY 2284 ATTGAAGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAAATCCAAACAAA 2343

Db 610 uLeuLysAsnGluGlnThrTrpArgAlaAspGluIleLeuProSerGluSerLysGlnLy 630  
QY 2344 GGATGATCAACAAAATTTCTTGGGATTTTGAGATTTCTTGGAGATCTCTTACAGATGA 2403  
Db 630 sAspTyrGluLysAsnSerTrpAspThrGluSerLeuCysGluThrValSerGlnLysAs 650  
QY 2404 TGTGTGTTTATCCCAAGCTACACATCAAAAAGATTCGATACCTTAAGTGGAAAATAGA 2463  
Db 650 pValCysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGl 670  
QY 2464 AGACTCTCTCATAAAGATGCTCTTCTGAAGCCTACCTGTGGAATGAAAATTTCTCTCC 2523  
Db 670 uGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePr 690  
QY 2524 AAAATAAGACCTTTAGAATTGAAGGACAGAGAAACATTCAAAAGCAGAGCATGTGAGTTCTGT 2583  
Db 690 oThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys-- 709  
QY 2584 AGACTCCACATTCAGTCTTTTGGCAACCCGACTACTGAAAATTCACAGTCTCAAAAAGT 2643  
Db 710 -ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal----- 723  
QY 2644 TGAGGAAGACTTTAATCTTACTACCAAGGAGGAGCAACAAAGACAGCTAACTGGACAACA 2703  
Db 724 724  
QY 2704 GGAACGTGATTTGGCATTATTGACGAGCTCCACAAGATCAAAACAAATAAGATGCCAC 2763  
Db 732 nGluGlnThrLeuArgAlaAspGluLeuLeuPro----- 743  
QY 2764 ATCAGATTTAGGAAGAAAGATACAAAATCAACTTCAGATCTCTGAGATTATCTCTGT 2823  
Db 744 -SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLysGly 763  
QY 2824 GAGTGATACACAGAAATTTATGAGTGTTCCTGAGGCTACATATCAAAAAGAAATAAGAC 2883  
Db 763 uThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIleAspLy 783  
QY 2884 AACAAATGGCAAAATAGAGAGTCT----- 2908  
Db 783 sIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCysAr 803  
QY 2908 2908  
Db 803 gMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAl 823  
QY 2909 -----CCTGAAAGCCTTCTCAGTTTGAGCTGCCACTGAAATGCAAACTCTGTTC 2961  
Db 823 aGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerValPr 843  
QY 2962 AAATAAGCCTTAGAATGGAGAAATAAACAAACATTCAGAGCA----- 3004  
Db 843 oAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPheProSe 863  
QY 3004 3004  
Db 863 rGluSerLysGlnLysLysValGluGluAsnSerTrpAspSerGluSerLeuArgGluTh 883  
QY 3004 3004  
Db 883 rValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLysI 903  
QY 3005 -----GATTCAACTACCTATCAAAAATCTTGGATGCACTCTCTCTTCTG 3048  
Db 903 eSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSerCy 923  
QY 3049 TGAAGAGGAGGAACTTAAAGAGATACTGTGAAACAAATTCACGCAAAAATGGAACA 3108  
Db 923 sGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGluGl 943  
QY 3109 AATGAAAAATAAGTTTGTGTACTACAAAGGAACCTGTGAGAGCGCAAGAAATATAATC 3168  
Db 943 nMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLysSe 963





Db 187 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 206  
QY 1979 TTCGATACCTTAAGTCGAAATTAAGAGAGTCTCTCTGTAAGAGCT 2038  
Db 207 IleAspLysIleAsnGlyLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 226  
QY 2039 ACTGTGAGGAAGATTCTCTTCCAAATAAAGCTTAGAATTAAAGGACAGAGAAACA 2098  
Db 227 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 246  
QY 2099 TTCAAAGCAGAGTCTCTGATAAAGATGCTCTCTGAAAGCTTACCTGTGGAAGAAAGTT 2158  
Db 247 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 266  
QY 2159 TCTCTCCAAATAAAGCCTTAGAATAAAGGACAGAGAAACAACACTCAAGCAGAGTCTCT 2218  
Db 267 SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 284  
QY 2219 GATAATGATGCTCTCTGAGCCTACCTGTGGAAGAAAGTTCTCTTCCAAATAAAGCT 2278  
Db 284  
QY 2279 TTAGAATTGAAGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAAATCCAAA 2338  
Db 285  
QY 2339 CAAAGGATGATGAAGAAATCTTGGGATTTTGAGAGTTTCCCTTGAGACTCTCTTACAG 2398  
Db 293 GlnLysAspTyrGluGluAsnSerTrpAspThrGluSerLeuCysGluThrValSerGln 312  
QY 2399 AATGATGCTGTGTTACCCAGGCTACATCAAAAAGAAATTCGATACCTTAAGTGGAAA 2458  
Db 313 LysAspValCysLeuProLysAlaLahisGlnLysGluIleAspLysIleAsnGlyLys 332  
QY 2459 TTAGAAGAGTCTCCTGATAAGATGGTCTTCTGAAAGCCTACTCTGGAATGAATAATTCT 2518  
Db 333 LeuGluGlySerProGlyLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSer 352  
QY 2519 CTTCCAAATAAAGCCTTAGATTGAAGCAGAGAAACATTCAAGCAGAGGAGTCTGAGT 2578  
Db 353 IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGlu 372  
QY 2579 TCTGAGAGTCCACATTCAGTCTTTTGGCAACCGACTACTGAAAATTCACAGTCTACA 2638  
Db 373 Lys---ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal 387  
QY 2639 AAAGTTGAGGAGACTTTAATCTTACTACAGGAGGAGGACACAAAGCAGTACTGGA 2698  
Db 388  
QY 2699 CAACAGGAACGTGATATTGGCATTATTGAACGAGCTCCACAAAGATCAACAAATAAGATG 2758  
Db 395 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro----- 407  
QY 2759 CCCACATCAGAAATTAGGAAGAAAGATACAAATCAATTCAGATTCGAGATTATC 2818  
Db 408 -----SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 425  
QY 2819 TCTGTGAGTGATACACAGAATTATGAGTGTCTTACCTGAGGCTACATATCAAAAAGAAATA 2878  
Db 426 CysGluThrValSerGlnLysAspValCysLeuProLysAlaLahisGlnLysGluIle 445  
QY 2879 AAGACAAATAAGGCAAAATGAAGAGTCT----- 2908  
Db 446 AspLysIleAsnGlyLysLeuGluGluSerProAsnAspGlyPheLeuLysSerPro 465  
QY 2908 ----- 2908  
Db 466 CysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 485  
QY 2909 -----CCTGAAAGCCTTCTCAGCTTTCAGCTGCGCTGCAATGCAAACTCT 2956

Db 486 LysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 505  
QY 2957 GTTCCCAATTAAGCTTTAGAAATGGAAGATAAACAACATTTGAGAGCA----- 3004  
Db 506 ValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGlnMetPhe 525  
QY 3004 ----- 3004  
Db 526 ProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSerLeuArg 545  
QY 3004 ----- 3004  
Db 546 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 565  
QY 3005 -----GATTCAACTACCTATCAAAAAATCTTGGATGCACTTCCT 3043  
Db 566 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 585  
QY 3044 TCTTGTCAAGAGAGGGAACCTTAAAGAAGATACTGTGAACAAATTTACAGCAAAATG 3103  
Db 586 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 605  
QY 3104 GAACAAATGAAAAATAAGTTTTGTACTACAAAAAGGAACCTGTCAAGAGCAAAAGAAATA 3163  
Db 606 GluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIle 625  
QY 3164 AATCAGAGTTAGAGAACCAAAAGCTAAATGGAACAGAGCTCTGCAGTGTGAGATTG 3223  
Db 626 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnLeuCysSerValA-gLeu 645  
QY 3224 CCTTTAAATCAAGAAGAAGAGAGAAATGTCGATATATTAAGAAAGAAAAATTAGA 3283  
Db 646 ThrLeuAsnGlnGluGluGluLysArgAsnAlaAspIleLeuAsnGluLysIleArg 665  
QY 3284 CCC-----GAAGAGCACTTAGAAAAAGTTAGAAAGTGAACCAACTT 3328  
Db 666 GluGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnLeu 685  
QY 3329 GAACAGACTCTCAGATACAGATATAGAAATTCGAAAGTCAACAAGTAAATTTGAATCAG 3388  
Db 686 GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln 705  
QY 3389 GTTCTCAGACTCATGAAAGTGAATGATCTCTTTTCATGAAAATTCATGTTGAAAAAG 3448  
Db 706 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 725  
QY 3449 GAATTTGCCATGCTAAACTGGAAGTAGCCACACTGAAACATCAACACAGGTGAAGGAA 3508  
Db 726 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGlu 745  
QY 3509 AATAAATACCTTTGAGGACATTAAGATTTTACAAAGAAAAAGAAATGCTGAACTTCAAAATGACC 3568  
Db 746 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 765  
QY 3569 CTAAACTGAAACAGAAACAGTAAACAAAAGGGCATCTCAGTATAGAGAGAGCTTAA 3628  
Db 766 LeuLysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 785  
QY 3629 GTTCTCAGCGCAGAGAACGATGCTGACTTCTTAAATTTGAAGGAA 3673  
Db 786 ValIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 800  
RESULT 10  
AAU33357  
ID AAU33357 standard; protein; 1095 AA.  
XX AC AAU33357;  
XX DT 18-DEC-2001 (first entry)  
XX DE Human breast cancer protein B726P fusion protein #1.  
XX KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;

KW gene therapy; immunogen.  
 XX Homo sapiens.  
 XX WO200179286-A2.  
 PN 25-OCT-2001.  
 XX 12-APR-2001; 2001WO-US012164.  
 XX 17-APR-2000; 2000US-00551621.  
 PR 08-JUN-2000; 2000US-00590751.  
 PR 22-JUN-2000; 2000US-00604287.  
 PR 20-JUL-2000; 2000US-00620405.  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;  
 PI WPI; 2001-611721/70.  
 XX N-PSDB; AAS47421.  
 DR Breast Tumor Proteins and nucleic acids useful for the prevention,  
 XX diagnosis and treatment of breast cancer.  
 PT  
 PT  
 XX Claim 22; Page 292-295; 297pp; English.  
 PS  
 XX The invention relates to isolated breast tumour proteins and nucleic  
 CC acids that encode them, including immunogenic fragments of the proteins.  
 CC Also included are expression vectors expressing the proteins, transformed  
 CC cells and antibodies raised against the proteins or an antigen presenting  
 CC cell expressing the protein. The proteins and nucleic acids may be used  
 CC in the prevention, diagnosis and treatment of diseases associated with  
 CC inappropriate breast tumour protein expression, i.e. breast tumours and  
 CC breast cancer e.g by gene therapy. The nucleic acids and their  
 CC complements may also be used as DNA probes in diagnostic assays to detect  
 CC and quantitate the presence of similar nucleic acids in samples, and  
 CC therefore which patients may be in need of restorative therapy. The  
 CC proteins, nucleic acids and antibodies may be used in assays to identify  
 CC modulators (e.g. antagonists) of breast tumour protein expression and  
 CC activity. The antibodies and antagonists may also be used to down  
 CC regulate expression and activity. The antibodies may also be used as  
 CC diagnostic agents for detecting the presence of the proteins in samples  
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-  
 CC purification diagnostic techniques. The present sequence is a breast  
 CC tumour protein encoded by a cDNA from a breast tumour cDNA library  
 CC isolated by subtractive hybridisation against a normal breast cDNA  
 CC library  
 XX  
 XX Sequence 1095 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1,18e-184 Length: 1095  
 Score: 2372.50 Matches: 529  
 Percent Similarity: 59.37% Conservative: 73  
 Best Local Similarity: 52.17% Mismatches: 151  
 Query Match: 36.79% Indels: 261  
 DB: 4 Gaps: 13  
 US-09-602-362E-26 (1-3673) x AAU33357 (1-1095)  
 QY 950 CTTTTCGACATATACGAAATTTACCTAAATAATCCATCCAGAGAACCA 1009  
 DB 24 LeuLeuGluAsnValIleSerLysThrIleAsnProGlnValSerLysThrGlu----- 41  
 QY 1010 TCTACAGGAACACCTGATGAGGTGCACCTTGGCGGAAGAACACCTGACACGCTGAA 1069  
 DB 42 -----TyrLys 43  
 QY 1070 AGCTTCTGGAATAAACACCTGACGAGGTGCACGCTTGGTGGAGGAACGCTTCTGCCAAA 1129  
 DB 44 GluLeuGlnGluPheIleAspAsnAla-----ThrThrAsnAla 58

QY 1130 ATTCAATGTCCTGGGAAAGCAACATCTGGAAGAGTTTGAACAGTCAACACAGAAACACCT 1189  
 DB 59 IleAspGluLeuLysGluCys-----PheLeuAsnGlnThrAspGluThrLeu 74  
 QY 1190 AGGAAATTT-----TTG 1201  
 DB 75 SerAsnValGluValPheMetGlnLeuIleTyrAspSerSerLeuCysAspLeuPheMet 94  
 QY 1202 AGCCTACAAAGAAACATCTGAGAAATTTTCATGCCAGCAAGAAAGAAAGATCTAGGAAG 1261  
 DB 95 SerProAlaLysGluThrSerGluLysPheThrTrpAlaAlaLysGlyArgProArgLys 114  
 QY 1262 ATCAGATGGGAGGAAAGAAACATCTGTAAGACATGTAATGCGTGGCAGGATTAACACCT 1321  
 DB 115 IleAlaTrpGluLysGluThrProValLysThrGlyCysValAlaArgValThrSer 134  
 QY 1322 AATAAACTGAAGTTTGGAAAAAGGAACATCTAATATGATTGCAATGCTCTCAAAAGAA 1381  
 DB 135 AsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGlu 154  
 QY 1382 ACATCTACAAAGCAAGTACAAATGTGGATGTGAGTCTGTAGAGCCTATATTCAGTCTT 1441  
 DB 155 SerSerThrLysAlaSerAlaAsn----- 162  
 QY 1442 TTTGGCACCGGACTATTGAAATTCACAGTGTACAAAGTTTTCAGGAGACTTTAATCTT 1501  
 DB 162 ----- 162  
 QY 1502 GCTACCAAGATTATCTTAAGAGTGTGCACAGAATTATACGTGTACGTGATGCTACA 1561  
 DB 162 ----- 162  
 QY 1562 TATCAAAAGATATCAAAACAAATAATCAAAATAGAAGATCAGATGTTCCCATCAGAA 1621  
 DB 163 -----AspGlnArgPheProSerGlu 169  
 QY 1622 TCCAAACGAGAGGAAGTAAAGAAATATCTTGGATCTCGGAGTCTCTTTGAGAGTTCT 1681  
 DB 170 SerLysGlnGluAspGluGluTyrSerCysAspSerArgSerLeuPheGluSerSer 189  
 QY 1682 GCAAGACTCAAGTGTATACCTGATCTATCTATCAGAAAGTAAATGAGATTAATAGA 1741  
 DB 190 AlaLysIleGlnValCysIleProGluSerIleTyrGlnLysValMetGluIleAsnArg 209  
 QY 1742 GAAGTAGAAGAGTCTCTGAGAGCCATCTGCTTCAAGCCTGCCGTGNGAAATGCAAAAG 1801  
 DB 210 GluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGlnAsn 229  
 QY 1802 ACTGTTCCAAATAAGCCTTTGAATTGAAGATGAACAAACATTTGAGACGCTCAGATG 1861  
 DB 230 SerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspProMet 249  
 QY 1862 TTCCCATCAGATCCAAACAAAGAGCAGATGAAGAAAAATCTTTGGGATCTGAGAGTCCC 1921  
 DB 250 PheProGluSerLysGlnLysAspTyrGluAsnSerTrpAspSerGluSerLeu 269  
 QY 1922 TGTGAGACGGTTTCACAGAGAGTGTATTTATCCCAAGCTACACATCAAAAGAAATTC 1981  
 DB 270 CysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluIle 289  
 QY 1982 GATACCTTTAGTGGAAAAATTAGAAGAGTCTCTGTTTAAAGATGCTCTTCTGAAGCCTACC 2041  
 DB 290 AspLysIleAsnGlnLysLysGluGluSerProAsnLysAspGlyLeuLeuLysAlaThr 309  
 QY 2042 TGTGGAAGGAAAGTTCTCTTCCAAATAAGCCTTAGAATTAAGGACAGAGAAACATTC 2101  
 DB 310 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThrPhe 329  
 QY 2102 AAAGCAGGTCTCTGATAAAGATGGTCTTCTCAAGCCTACCTGTGTGGAAGAAAGTTTCT 2161  
 DB 330 LysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLysSer 349  
 QY 2162 CTTCCAAATAAGCCTTAGAATTAAGGACAGAGAAACACTCAAGACAGAGTCTCTGAT 2221



PR	22-JUN-2000; 2000US-00604287.
PR	20-JUL-2000; 2000US-00620405.
XX	(CORI-) CORIXA CORP.
XX	Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI	Henderson RA;
XX	WPI; 2002-635657/68.
DR	N-PSDB; ABS64022.
XX	
PT	Novel breast cancer polynucleotides and polypeptides encoded by the
PT	polynucleotides, useful for detecting the presence of breast cancer in a
PT	patient, and in pharmaceutical compositions, for treating breast cancer.
XX	
PS	Disclosure; Page 223-225; 247pp; English.
XX	
CC	The invention relates to an isolated breast tumour polynucleotide and the
CC	polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC	detecting the presence of breast cancer in a patient, and in
CC	pharmaceutical compositions for treating breast cancer. The sequences are
CC	useful for stimulating an immune response in a patient and can therefore
CC	be used in production of vaccines. The sequences are also useful for
CC	detecting the presence of a cancer in a patient, by obtaining a
CC	biological sample from the patient, contacting the biological sample with
CC	a composition of the invention and detecting the amount of polynucleotide
CC	that hybridizes to the sample. This sequence represents a human breast
CC	tumour polypeptide of the invention
XX	
SQ	Sequence 1095 AA;
	Alignment Scores:
Pred. No.:	1.18e-184 Length: 1095
Score:	2372.50 Matches: 529
Percent Similarity:	59.37% Conservative: 73
Best Local Similarity:	52.17% Mismatches: 151
Query Match:	36.79% Gaps: 261
DB:	5 TyLys 13
US-09-602-362E-26 (1-3673) x ABG78924 (1-1095)	
QY	950 CTTTGGACATATACGAAATTACCTATAAAATCCTCAAATATCCAATCAGAGGAACA 1009
Db	LeuLeuGlunValIleSerLysThrIleAsnProGlnValSerLysThrGlu----- 41
QY	1010 TCTACGAACACTGTATGAGCGTGCACTTGGCGGAAGAACAACCTGACAGCGGTAA 1069
Db	-----TyLys 43
QY	1070 AGCTTGCTGGAAAAAACACCTGACAGGCTGCACGCTTGGTGGAGGAACTCTGCCAAA 1129
Db	GlulLeuLeuGlnGluPheIleAspAsnAla-----ThrThrAsnAla 58
QY	1130 ATTCAATGTCCTGGGAAAGCAACATCTGGAAGTTTGAAACAGTCAACAGAAGAAACACCT 1189
Db	IleAspGluLeuLysGluCys-----PheLeuAsnGlnThrAspGluThrLeu 74
QY	1190 AGGAAATT-----TTG 1201
Db	SerAsnValGluValPheMetGlnLeuIleTyRaspSerSerLeuCysAspLeuPheMet 94
QY	1202 AGGCCTACAAAAGAAACATCTGAGAAATTTTCATGTGGCCAGCAAAAGAAAGATCTAGNAG 1261
Db	SerProAlaLysGluThrSerGluLysPheThrTrpAlaLysGlyArgProArgLys 114
QY	1262 ATCACATGGGAGGAAAAAGAAACATCTGTAAGACTCAATGCGTGGCAGGATTAACACCT 1321
Db	IleAlaTrpGluLysGluThrProValLysThrGlyCysValAlaArgValThrSer 134
QY	1322 AATAAACTGAAGTTTTGGAAAAAGGAACATCTAATATGATTCATGCTCTCAACAAGAA 1381
Db	AsnLysThrLysValLeuGluLysGlyArgSerLysMetIleAlaCysProThrLysGlu 154

Db 416 GluGlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIle 435  
 QY 2522 CCAATTAAGCCCTTGAATTTGAAGACACAGAGAAACATTCAAGCAGAGAGATGTGAGTTCT 2581  
 Db 436 ProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLys 455  
 QY 2582 GTAGAGTCCACATTGAGTCTTTTGGCAACCGACTACTGTGAAATTCACAGTCTACAAA 2641  
 Db 456 ---ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal--- 469  
 QY 2642 GTTAGGAGAGACTTTAACTTACTACCAAGGAGGAGCAACAAGACAGTACTGGACAA 2701  
 Db 470 -----ProAsnLysAlaLeuGluLys 477  
 QY 2702 CAGGAACGTGATTTGGCATTATTGAACGAGCTCCCAAGATCAAAACAATAAGATGCC 2761  
 Db 478 AsnGluGlnThrLeuArgAlaAspGluIleLeuPro----- 489  
 QY 2762 ACATCAGAAATTAGGAAGAAAGAAAGATCAAAATCAACTTCAGATTCTGAGATTATCTCT 2821  
 Db 490 ---SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 508  
 QY 2822 GTGAGTGATACACAGAAATTATGAGTGTTCCTGAGGCTACATATCAAAAAGAAATAAG 2881  
 Db 509 GluThrValSerGlnLysAspValCysLeuProLysAla\*\*HisGlnLysGluIleAsp 528  
 QY 2882 ACACAAAATGGCAAAATAGAGAGTCT----- 2908  
 Db 529 LysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaProCys 548  
 QY 2908 ----- 2908  
 Db 549 ArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 568  
 QY 2909 -----CCTGAAAGCCTTCTCATTGAGCCTGCCTGAAATTCGAAATCTCTCTT 2959  
 Db 569 AlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 588  
 QY 2960 CCAATTAAGGCTTGAATGGAAGATAAACAACATTGAGAGCA----- 3004  
 Db 589 ProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGlnMetPhePro 608  
 QY 3004 ----- 3004  
 Db 609 SerGluSerLysGlnLys\*\*ValGluGluAsnSerTrpAspSerGluSerLeuArgGlu 628  
 QY 3004 ----- 3004  
 Db 629 ThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAspLys 648  
 QY 3005 -----GATTCACTACCTTATCAAAAATCTTGATGCACTTCTTCT 3046  
 Db 649 IleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHisSer 668  
 QY 3047 TGTGAAAGAGGAGGAACTTAAAGATAACTGTGAACAAATTCACAGCAAAATGAA 3106  
 Db 669 CysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMetGlu 688  
 QY 3107 CAAATGAAATAAGTTTGTCTACTACAAAGCAACTGTCAAGCCGAAAGAAATAAAA 3166  
 Db 689 GlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLysGluIleLys 708  
 QY 3167 TCACAGTTAGAGAACCAAAAGCTAAATGGGAACCAAGAGCTCTGAGTGTGAGATTGCT 3226  
 Db 709 SerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerValArgLeuThr 728  
 QY 3227 TTAATCAAGAGAGAGAGAGAGAAATGTGATATATTAAAGAAAAATAGACCC 3286  
 Db 729 LeuAsnGlnGluGluLysArgAsnAlaAspIleLeuAsnGluLysIleArgGlu 748  
 QY 3287 -----GAGAGCAACTTAGGAAAGATTAGAGTCAACACCAACTTCAA 3331

Db 749 GluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeuGlu 768  
 QY 3332 CAGACTCTCAGATACAGATATAGAAATTTGAAAAGTGTAAACAAGTAAATTTGAATCAGGTT 3391  
 Db 769 GluAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGlnVal 788  
 QY 3392 TCTCACACTCATGAAAGTGAATGATCTCTTTTCATGAAAATTCGATGTTGAAAAAGGAA 3451  
 Db 789 SerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLysGlu 808  
 QY 3452 ATTGCCATGCTAAACATGGAGTAGCCACACTCAAAATCAACACACAGGTGAAGGAAAT 3511  
 Db 809 IleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnIleGlnGluLysGluAsn 828  
 QY 3512 AAATACTTTGAGGACATTAAGATTTTACAGAAAAGAAATGCTCAACTTCAAAATGACCTTA 3571  
 Db 829 LysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeu 848  
 QY 3572 AAATGNAACAGAAACAGTAAACAAGAGGCGCATCTCAGTATAGAGAGCAGCTTAAAGTT 3631  
 Db 849 LysLeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysVal 868  
 QY 3632 CTGACGCGACAGAACACGATGCTGACTTTAAATTTGAAGGAA 3673  
 Db 869 LeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 882  
 RESULT 12  
 ABJ37747  
 ID ABJ37747 standard; protein; 1095 AA.  
 XX AC ABJ37747;  
 XX DT 15-MAY-2003 (first entry)  
 DE Human tumour-related protein - SEQ ID No 493.  
 KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;  
 KW tumour; breast cancer; cancer; immune response stimulation.  
 OS Homo sapiens.  
 XX WO200283956-A1.  
 XX PD 24-OCT-2002.  
 XX PF 15-APR-2002; 2002WO-US012378.  
 XX PR 13-APR-2001; 2001US-00834759.  
 PR 07-DEC-2001; 2001US-00007805.  
 PR 13-FEB-2002; 2002US-00076622.  
 XX (CORI-) CORIXA CORP.  
 PA Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;  
 PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;  
 PI Vedvick TS, McNeill PD, Durham M;  
 XX WPI; 2003-103376/09.  
 PT New polypeptide and polynucleotide useful for stimulating and/or  
 FT expanding T cells specific for a tumor protein and treating breast  
 XX cancer.  
 PS Disclosure; Page 310-312; 375pp; English.  
 CC The invention comprises a method of stimulating and/or expanding T cells  
 CC specific for a tumour protein. The invention further comprises human  
 CC nucleic acids and proteins that are associated with tumours (e.g. breast  
 CC cancer). The method and sequences of the invention are useful for  
 CC stimulating and/or expanding T cells specific for a tumour protein,  
 CC detecting the presence of cancer, stimulating an immune response in a  
 CC patient and treating breast cancer. The present amino acid sequence  
 CC represents a human tumour-related protein











QY 1859 ATGTTCCCATCAGAAATCCAAACAAAGGACGATGAAGAAATTTCTTGGGATCTTGAGACT 1918  
Db MetPheProGluSerLysGlnLysAspTyrGluGluAsnSerTrpAspSerGluSer 175  
QY 1919 CCCTGTGAGAGGTTTACAGAGAGTGTGATTTTACCCAAAGCTACACATCAAAAGAA 1978  
Db LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195  
QY 1979 TTCGATACCTTAAGTGGAAATTTAGAGAGCTCTCTGTTAAAGATGGTCTTCTGAGCCT 2038  
Db IleAspLysIleAsnGlnLysLeuGluGluSerProAsnLysAspGlyLeuLeuLysAla 215  
QY 2039 ACCTGTGGAAGAAAGTTTCTCTTCCAAATAAAGCCTTAGAATTAAGACAGACAGAAACA 2098  
Db ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuLysAspMetGlnThr 235  
QY 2099 TTCAAAGCAGAGCTCTCTGATTAAGATGTTCTTCTGAAGCTACCTGTGGAAGAAAGTT 2158  
Db PheLysAlaGluProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255  
QY 2159 TCTCTTCCAAATTAAGCCTTAGAATTAAGGACAGAGAAACACTCAAGCAGAGCTCTCT 2218  
Db SerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 273  
QY 2219 GATAATGATGTTCTTCTGAAGCCTACTCTGTGAAGGAAAGTTTCTTCTTCCAAATAAAGCT 2278  
Db 273  
QY 2279 TTAGAATTGAAGGACAGAGAAACATTCAAAGCGCTCAGATGTTCCCATCAGATCCAAA 2338  
Db GluLeuProSerGluSerLys 281  
QY 2339 CAAAGCATGATGAAGAAATTTCTGGGATTTTGAGAGTTTCTCTGAGACTCTCTACAG 2398  
Db GlnLysAspTyrGluGluSerTrpAspSerGluSerLeuLysGluThrValSerGln 301  
QY 2399 AATGATGCTGTGTACCCAGCTACACATCAAAAGAAATTCGATACCTTAAGTGGAAA 2458  
Db LysAspValCysLeuProLysAlaHisGlnLysGluLeuAspLysIleAsnGlyLys 321  
QY 2459 TTAGAAGAGTCTCTGATTAAGATGTTCTTCAAGCCTACTCTGGAATGAAATTTCT 2518  
Db LeuGluLysSerProValLysAspGlyLeuLysAlaAsnCysGlyMetLysValSer 341  
QY 2519 CTTCGAATTAAGCCTTAGAATTAAGGACAGAGAAACATTCAAAGCAGAGATGTGAGT 2578  
Db IleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProGlu 361  
QY 2579 TCTGTAGAGTCCACATTCAGTCTTTTGGCAACCGACTACTGAAATTCAGTCTTACA 2638  
Db 362 Lys---ProSerAlaPhe-----GluProAlaIleGluMetGlnLysSerVal 376  
QY 2639 AAGTTGAGAAAGCTTTAATCTTACTACAGGAGGAGCAACAAAGACAGTAACTGGA 2698  
Db 377 -----ProAsnLysAlaLeuGluLeu 383  
QY 2699 CAACAGGACGATGATTTGGCATTATTGAACGAGCTCCACAGATCAACAAATAAGATG 2758  
Db 384 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuPro----- 396  
QY 2759 CCCATCATCAGAAATTAAGGAAAGAGATCAAAATCAACTTCAGATTCTGAGATTATC 2818  
Db 397 -----SerGluSerLysGlnLysAspTyrGluGluSerTrpAspSerGluSerLeu 414  
QY 2819 TCTGTGAGTGTACACAGATTAAGTGTGTACCTGAGCTGATCATATCAAAAGAAATA 2878  
Db 415 CysGluThrValSerGlnLysAspValCysLeuProLysAlaHisGlnLysGluIle 434  
QY 2879 AAGACAAACAAATGGCAAAATAGAGAGTCT----- 2908  
Db 435 AspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLysAlaPro 454

QY 2908 ----- 2908  
Db 455 CysArgMetLysValSerIleProThrLysAlaLeuLeuMetAspMetGlnThrPhe 474  
QY 2909 -----CCTGAAAAGCCTTCTCAGCTTTGAGCTGCCCTGAAATGCAAACTCT 2956  
Db 475 LysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSer 494  
QY 2957 GTTCCAAATAAGCCTTAGAATGGAAGAAATAAACAAACATTTGAGACA----- 3004  
Db 495 ValProAsnLysAlaLeuGluLeuLysAsnGlnThrLeuArgAlaAspGlnMetPhe 514  
QY 3004 ----- 3004  
Db 515 ProSerGluSerLysGlnLys\*\*\*ValGluGluAsnSerTrpAspSerGluSerLeuArg 534  
QY 3004 ----- 3004  
Db 535 GluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGluMetAsp 554  
QY 3005 -----GATTCAACTACCTATCAAAAATCTTGGATGCACCTCTCT 3043  
Db 555 LysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThrValHis 574  
QY 3044 TCTTGTGAAAGAGGAGGAACTTTAAAGAAAGATACTGTGAACAAATTCAGCAAAATG 3103  
Db 575 SerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnArgThrGlyLysMet 594  
QY 3104 GAACAAATGAAATTAAGTTTGTGTACTACAAAGGAACTGTGCAAGCGAAGAAATA 3163  
Db 595 GluGlnMetLysLysLysPheCysValLeuLysLysLeuSerGluAlaLysGluIle 614  
QY 3164 AAATCAGATTTAGAGAACCAAAAAGCTAAATGGGAAACAGAGCTCTGCAGTGTGAGATTG 3223  
Db 615 LysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnLysCysSerValArgLeu 634  
QY 3224 CCTTTAATCAAGAGAGAGAGAGAAATGTCGATATATTAAGAAAGAAATTAAG 3283  
Db 635 ThrLeuAsnGlnGluGluLysArgAsnAlaAspIleLeuAsnGluLysIleArg 654  
QY 3284 CCC-----GAAGACCACTTAGGAAAAGTTAGAACTGAAACACCAACTT 3328  
Db 655 GluGluLeuGlyArgIleGluGluGlnHisArgLysGluLeuGluValLysGlnGlnLeu 674  
QY 3329 GACAGACTCTCAGATACAGATATAGAAATGAAAAGTGAACAGTAACTGAATCAG 3388  
Db 675 GluGlnAlaLeuArgIleGlnAspIleGluLeuLysSerValGluSerAsnLeuAsnGln 694  
QY 3389 GTTCTTCACACTCATGAAAGTGAATGATCTCTTTCATGAAAATGTCATGTTGAAAAAG 3448  
Db 695 ValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCysMetLeuLysLys 714  
QY 3449 GAAATTCATGCTGATAAAGTGAAGTAGCCACACTGAAACATCAACACAGGTGAAGAA 3508  
Db 715 GluIleAlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGlu 734  
QY 3509 AATAAATACTTTCAGACATTAAGATTTTACAGAAAGAAAGTGAACCTCAATGACC 3568  
Db 735 AsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThr 754  
QY 3569 CTAAGACTGAACAGAAAAACAGTAACAAAAGGGCATCTCAGTATAGAGAGCAGCTTAA 3628  
Db 755 LeuLysLeuLysGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLys 774  
QY 3629 GTTCTCAGCGCAGAGACAGATGCTGACTTCTAAATTCAGAGAA 3673  
Db 775 ValLeuIleAlaGluAsnThrMetLeuThrSerLysLeuLysGlu 789

RESULT 15

ABJ37741

ID ABJ37741 standard; protein; 1002 AA.

XX

AC ABJ37741;

XX 15-MAY-2003 (first entry)  
 XX Human tumour-related protein - SEQ ID No 475.  
 DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;  
 KW tumour; breast cancer; cancer; immune response stimulation.  
 XX Homo sapiens.  
 OS  
 XX WO200283956-A1.  
 PN 24-OCT-2002.  
 XX 15-APR-2002; 2002WO-US012378.  
 XX 13-APR-2001; 2001US-00834759.  
 PR 07-DEC-2001; 2001US-00007805.  
 PR 13-FEB-2002; 2002US-00076622.  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;  
 PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;  
 PI Vedwick TS, McNeill PD, Durham M;  
 XX WPI; 2003-103376/09.  
 DR  
 XX New polypeptide and polynucleotide useful for stimulating and/or  
 PT expanding T cells specific for a tumor protein and treating breast  
 PT cancer.  
 XX  
 PS Example 1; Page 298-300; 375pp; English.  
 CC The invention comprises a method of stimulating and/or expanding T cells  
 CC specific for a tumour protein. The invention further comprises human  
 CC nucleic acids and proteins that are associated with tumours (e.g. breast  
 CC cancer). The method and sequences of the invention are useful for  
 CC stimulating and/or expanding T cells specific for a tumour protein,  
 CC detecting the presence of cancer, stimulating an immune response in a  
 CC patient and treating breast cancer. The present amino acid sequence  
 CC represents a human tumour-related protein  
 XX  
 XX Sequence 1002 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 3.19e-184 Length: 1002  
 Score: 2367.00 Matches: 511  
 Percent Similarity: 62.62% Conservative: 62  
 Best Local Similarity: 55.85% Mismatches: 126  
 Query Match: 36.71% Indels: 216  
 DB: 6 Gaps: 9  
 US-09-602-362E-26 (1-3673) x ABU37741 (1-1002)  
 QY 1199 TTGAGCCCTACAAAAGAAACATCTGAGAAATTTTCATGCCGCAAGAAAGATCTAGG 1258  
 Db 1 MetSerProAlaLysGluThrSerGluLysPheThrTyrAlaLysGlyArgProArg 20  
 QY 1259 AAGATCACATCGGAGGAAAGAAACATCTGTAAGACTGAATCGCTGCAGAGTAACA 1318  
 Db 21 LysileAlaTyrGluLysLysGluThrProValLysThrGlyCysValAlaArgValThr 40  
 QY 1319 CCTAATAAATCTGAAGTTTGGAAAAGGAACATCTAATATGATTGCTGCTACAAA 1378  
 Db 41 SerAsnLysThrLysValLysGlyArgSerLysMetLysCysProThrLys 60  
 QY 1379 GAAACATCTACAAAGCAAGTACAAATGTGGATGTGAGTTCTGTAGAGCCTATATTCAGT 1438  
 Db 61 GluSerSerThrLysAlaSerAlaAsn  
 QY 1439 CTTTGGCCACCGACTATGAAAATTCACAGTGTACAAAAGTTGAGGAGACTTTAAT 1498

Db 69 ----- 69  
 QY 1499 CTTGTACCAAGATTATCTTAAGAGTGTGTGACAGAATTATACGTGTTTACTGATGCT 1558  
 Db 69 ----- 69  
 QY 1559 ACATATCAAAAAGATATCAAAAATAAATCAAAAATAGAAGATCAGATGTTCCCATCA 1618  
 Db 70 -----AspGlnArgPheProSer 75  
 QY 1619 GAATCCAAACGAGGAGGAGATGAAGAATATTTCTGGGAGTCTCTTTTGAGAGT 1678  
 Db 76 GluSerLysGlnGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 95  
 QY 1679 TCTGCAAGACTCAAGTGTGTATACCTGATCTATCTATCAGAAAGTAAATGAGGATAAAT 1738  
 Db 96 SerAlaLysIleGlnValCysIleProGluSerIleGlnLysValMetGluLysAsn 115  
 QY 1739 AGAAGAGTGAAGAGCTTCTGAGAGAGCCATCTGCTTCAAGCCCTCCCGTNGAAATGCAA 1798  
 Db 116 ArgGluValGluGluProProLysLysProSerAlaPheLysProAlaIleGluMetGln 135  
 QY 1799 AAGACTGTTCCAAATAAAGCTTTGAATTGAAGTAAGAACAAACATGAGAGAGCTCAG 1858  
 Db 136 AsnSerValProAsnLysAlaPheGluLeuLysAsnGluGlnThrLeuArgAlaAspPro 155  
 QY 1859 ATGTTCCCATCAGATCCAAACAAAGGAGCGATCAAGAAATTTCTGGGATTTCTGAGAT 1918  
 Db 156 MetPheProGluLysLysGlnLysAspTyrGluLysAsnSerTyrPaspSerGluSer 175  
 QY 1919 CCCTGTGAGAGCGTTTCACAGAAGGATGTATTTACCCAAAGCTACACATCAAAAAGAA 1978  
 Db 176 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 195  
 QY 1979 TTCGATACCTTAAGTGAATAATAGAGAGTCTCTGTTAAAGATGCTCTTCTGAAAGCT 2038  
 Db 196 IleAspLysIleAsnGlyLysLeuLysGluSerProAsnLysAspGlyLeuLysAla 215  
 QY 2039 ACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCTTGAATTAAGAGCAGAGAAACA 2098  
 Db 216 ThrCysGlyMetLysValSerIleProThrLysAlaLeuGluLysAspMetGlnThr 235  
 QY 2099 TTCAAGCAGAGTCTCTGATTAAGATGCTCTTCTGAGGCTACCTCTGGAAGAAAGTT 2158  
 Db 236 PheLysAlaGluProProGlyLysProSerAlaPheGluProAlaThrGluMetGlnLys 255  
 QY 2159 TCTCTTCCAAATAAAGCTTTAGAAATTAAGCAGAGAAACACCTCAAGCAGAGATCTCCT 2218  
 Db 256 SerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAsp----- 273  
 QY 2219 GATAATGATGCTCTTCTGAAAGCTTACCTGTGGAAGGAAAGTTTCTCTTCCAAATAAAGCT 2278  
 Db 273 ----- 273  
 QY 2279 TTGAATTTGAGGACAGAGAAACATTCAAAGCAGCTCAGATGTTCCCATCAGAAATCAAA 2338  
 Db 274 -----GluIleLeuProSerGluSerLys 281  
 QY 2339 CAAAAGGATGATGAAGAAATTTCTGGGATTTTGGAGATTTCTTGAGACTCTCTTTACAG 2398  
 Db 282 GlnLysAspTyrGluLysSerSerTyrAspSerGluSerLeuCysGluThrValSerGln 301  
 QY 2399 AATGATGTGTGTACCCCAAGCTACATCAAAAAGNATTCGATACCTTAAGTGGAAAA 2458  
 Db 302 LysAspValCysLeuProLysAla\*\*\*HisGlnLysGluIleAspLysIleAsnGlyLys 321  
 QY 2459 TTGAAGAGTCTCTGATGAAGAGATGCTCTTCTGAGGCTTACCTGTGGAAATGAAATTTCT 2518  
 Db 322 LeuGluLysSerProValLysAspGlyLeuLysAlaAsnCysGlyMetLysValSer 341  
 QY 2519 CTTCCAAATAAAGCTTTAGAAATTTGAAGCAGAGAAACATTCAAAGCAGAGAGTGTGAGT 2578  
 Db 342 IleProThrLysAlaLeuGluMetAspMetGlnThrPheLysAlaGluProProGlu 361

